Testing and Characterization of Selective Trans-Acting Hammerhead Ribozymes that Cleave Two Disease-Causing Mutant Transcripts of the PABPN1 Gene

Pegah Hadavi

A Thesis
in the Department of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of
Master of Applied Science in Electrical and Computer Engineering,
Concordia University,
Montreal, Quebec, Canada

December 2022

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Pegah Hadavi
Entitled: Testing and Characterization of Selective Trans-Acting Hammerhead Ribozymes that Cleave Two Disease-Causing Mutant Transcripts of the PABPN1 Gene

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Electrical and Computer Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. S. Shih

_____ External Examiner
Dr. M. Whiteway

_____ Internal Examiner
Dr. S. Shih

_____ Supervisor
Dr. N. Kharma

Approved by: _____
Dr. Yousef R. Shayan

Department of Electrical and Computer Engineering

March 2023

_____
Dr. Mourad Debbabi
Faculty of Engineering and Computer Science

Abstract

Testing and Characterization of Selective Trans-Acting Hammerhead Ribozymes that Cleave Two Disease-Causing Mutant Transcripts of the PABPN1 Gene

Pegah Hadavi

Ribozymes are RNA molecules with catalytic functions. These molecules can catalyze a self-cleaving reaction or bind to a target RNA molecule and cleave it. These properties make ribozymes perfect candidates for RNA therapeutics. Hammerhead ribozymes are a well-studied family of ribozymes and can be used to cleave specific mRNA molecules. These synthetically generated ribozymes can then regulate gene expression within cells. Normal PABPN1 protein has 10 alanine (GCG codon) repeats in its structure. Oculopharyngeal Muscular Dystrophy (OPMD), a hereditary disease with no cure, is caused by mutations that result in an increase in the number of GCG repeats in PABPN1 gene. The mutant proteins translated from genes that contain these repeat expansions are believed to be the disease-causing agents, hence targeting the expression of these proteins on the mRNA level is an attractive strategy using RNA therapeutics in OPMD treatment. To generate these RNA molecules, an evolutionary algorithm (*Tri-Cleaver*) was used to design ribozymes to be tested against two mutants of PABPN1 transcripts. Twenty-nine ribozymes that were designed using this algorithm were tested in human cells (HEK293) to investigate their effect on PABPN1 mutant transcripts with 13 and 17 alanine codon (GCG) repeats. These ribozymes were tested for 1) their efficiency to bind and cleave PABPN1 mRNA with 13 and 17 GCG repeats and 2) their selectivity for the mutant transcripts of PABPN1 gene. The results not only show the successful use of an evolutionary algorithm in designing trans-cleaving ribozymes that can selectively target repeat expansions, but also provide useful data to improve the algorithm's later designs.

Acknowledgements

I would like to extend my heartfelt gratitude to Dr. Nawwaf Kharma and Dr. Jonathan Perreault, for their extensive and professional guidance, motivation, and patience. I would like to express my deep gratitude to my mentor Dr. Aida Abu-Baker who supervised my work and guided me through my graduate research at McGill University. I must also acknowledge Mr. Nicolas Kamel's (PhD candidate) contribution to my thesis through the design, using his *TriCleaver* algorithm, of the selective hammerhead ribozymes that I tested. My grateful thanks are extended to Dr. Guy Rouleau who provided me with an opportunity to work in his laboratory and Daniel Rochefort who taught me many new lab skills. Finally, I would like to thank my partner Dave and my family for their unconditional support and encouragement through my studies.

Table of Contents

List of Figures

List of Tables

Abbreviations

10-Ala: 10 alanine repeats
13-Ala: 13 alanine repeats
17-Ala: 17 alanine repeats
cDNA: Complementary DNA
CT value: Cycle threshold value
DNA: Deoxyribonucleic acid
GCG: Guanine, Cytosine, Guanine nucleotide sequence which codes for the amino acid alanine
GFP: green fluorescent protein
HEK293: Human embryonic kidney 293 cells
hhRz: Hammer head ribozyme
kD: Kilodalton
M: molar
M13: Mutant PABPN1 gene with 13 alanine repeats
M17: Mutant PABPN1 gene with 17 alanine repeats
µg: microgram
µl: microliter
miRNA: MicroRNA
mRNA: Messenger RNA
OBS: Oligonucleotide binding sites
OPMD: Oculopharyngeal muscular dystrophy
PABPN1: Poly (A) binding protein, nuclear 1
Rbz: Ribozyme
RiBS: Ribozyme binding sites
RNA: Ribonucleic acid
RQ: relative quantification
RT-PCR: Reverse transcription-polymerase chain reaction
RT-qPCR: Quantitative reverse transcription-polymerase chain reaction
SDS-PAGE: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
shRNA: Small hairpin RNA
SUB: sodium dodecyl sulfate utilizing buffer
3'-UTR: 3' untranslated region
WT: Wild-type PABPN1 gene
HD: Huntington disease
MJD: Machado-Joseph disease
STR: Short tandem repeats
FRDA: Friedreich ataxia
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
EA: Evolutionary algorithm

Chapter 1: Introduction

1.1 Overview trinucleotide-repeat expansion disorders

Nucleotide repeat expansion disorders include over 40 diseases some of which occus within the same families with similar symptoms and complications [1, 2]. These disorders, which mainly cause neurological complications, are most frequently caused by trinucleotide repeat expansions, although larger number of nucleotide repeats can also be involved in the expansion (tetra-, penta-, hexa-, dodeca-) [3]. Trinucleotide-repeat expansion disorders are neuropsychiatric, and the age of onset of the disease is inversely proportional to the number of trinucleotide repeats exceeding the normal number of repeats. Examples of such diseases are Huntington disease (HD), Machado-Joseph disease (MJD), both cause by expansion in CAG (glutamate) repeats, and Oculopharyngeal muscular dystrophy (OPMD) caused by expansion of the GCG (alanine) repeats [4]. The normal number of repeats that will produce the healthy protein is 26 glutamate repeats for HD, 14 to 40 glutamate repeats for MJD, and 10 alanine repeats for OPMD [4, 5]. Studies suggest that crossing the normal threshold in the number of trinucleotide repeats does not immediately cause symptoms, but can mean that the individual is a carrier, and his/her children can be affected, as these repeats can increase in number from one generation to the next. The exact mechanism of these microsatellite expansions for each disease has yet to be determined, as these expansions can happen in different cells and at different stages of human development. These mechanisms may be disease specific or depend on the number of the repeats, and theses expansions have been mostly studied in organisms other than humans and may not be directly applicable to humans [4, 6]. Two of the proposed mechanisms by which these deleterious repeats are expanded are DNA slippage, and there is evidence that connects these expansions to insertion of transposable elements [6].

DNA slippage, which can occur during cell replication, gene conversion, DNA repair, and mismatch repair, has been shown to result in expansion of trinucleotides (or possibly larger number of nucleotides) [7]. DNA slippage happens when the DNA polymerase stops and falls off a strand of DNA that is being duplicated, modified or repaired. This slippage can result in addition or deletion of a single nucleotide or larger sequences of microsatellite DNA (short tandem repeats (STR)). When the DNA strands, which contain repeats separate, the single strands can form loops which can be stabilized by base pairs depending on the nucleotides (A, T, C, G) in the repeat. In case of GCG (alanine) repeats for example, the C.G or G.C base pair can cause these loops to form. When these loops form during replication or via other mechanisms that involve the DNA polymerase, they can cause a displacement (slip) in the strand, which means that when the DNA polymerase synthesizes the complementary strand to these repeats, these repeats can be expanded [4, 8]. The probability of DNA slippage may increase with the number of repeats, as the longer repeats can result in more stable hairpin loops.

Roughly 45% of the human genome is found to be repetitive sequences derived from transposable elements ("jumping genes") [1, 9]. Alu elements are transposable elements, which are the most abundant repeat sequences in the human genome (around 10%) [9]. These retrotransposable elements (transposable via RNA intermediates) are responsible for gene regulation through

different cellular mechanisms, such as regulation of translation of proteins, RNA editing and alternative splicing [10]. The movement of these elements in the genome starts when RNA polymerase III transcribes Alu element into RNA. This RNA forms a complex with proteins forming ribonucleoprotein (RNP), which is then reverse transcribed into a double stranded DNA, which will be inserted into a new location on a chromosome. The insertion of these elements into new locations in the genome is described as "jumping". When an Alu element jumps several copies of its DNA are produced. This means that the number of these elements in the human genome has been growing progressively as these primate-specific repeats have facilitated human evolution by regulating transcription and translation, creating new genes and transposable elements, and influencing gene expression involved in neurogenesis and the generation of the neural networks. Although Alu elements have contributed to human evolution, they are also responsible for several genetic disorders mainly because they cause genetic rearrangements that can be deleterious [11-13]. Alu insertions have been shown to be responsible for some nucleotide expansion disorders and can be the underlying genetic cause of others [14]. Inherited neurological disorders such as Friedreich ataxia (FRDA) have the GAA repeat expansions in the middle of Alu repeats showing that insertion of Alu elements may have a role in trinucleotide repeat expansion diseases [15]. Alu elements have also been shown to cause more frequent duplications in GC-rich part of the chromosome, which may suggest a role in increasing of the number of alanine repeats (GCN) in OPMD [16].

Regardless of the cause of the expansion, nucleotide repeats have been shown to be essential for genomic regulation [17, 18]. However, expansion of these repeats and particularly trinucleotide repeats can lead to various disorders. The number of these repeats can vary for each family of diseases and from one individual to another. In many nucleotide repeat expansion disorders, the expansion in the repeats began in parents and was passed down to offspring [6].

1.2 Oculopharyngeal Muscular Dystrophy (OPMD)

Oculopharyngeal muscular dystrophy (OPMD) is a progressive late-onset disease that is characterized by eyelid drooping, dysphagia (swallowing problems), facial and proximal muscle weakness, tongue shrinking, and diplopia (double vision) [19]. Studies show that the age of onset of the disease and the severity of the symptoms depend on the number of alanine repeats. The longer repeats are believed to be associated with earlier onset and more severe symptoms. The late onset normally starts on the 5th or 6th decade of life and although not considered lethal, OPMD can dramatically decrease the quality of patients' lives. The disease impairs the individuals affected by not only weakening the oculopharyngeal muscles, but also causing chronic pain and fatigue, progressive dementia, and loss of executive functions, respiratory issues, weakness in shoulder, hip and thigh muscles which reduces mobility gradually [20, 21]. In a study of a cohort of 89 OPMD patients age 37 to 84 (mean 66.2) , nearly half of the patients reported using some type of mobility assistive device such as walkers, canes, scooters, wheelchairs and about 95% of the patients suffered from dysphagia and ptosis (upper eyelid drooping) [22].

Muscle biopsies of patients show intranuclear inclusions which are known as the morphological indications of the disease [23]. Accumulation of these nuclear inclusions, which contain the misfolded PABPN1 protein along with other cellular components and nuclear proteins, interfere

with normal cell function and has been associated with cell death [24]. Abnormal expansion of GCG (or in some cases GCA) repeats in exon 1 of the poly (A) binding protein, nuclear 1 (PABPN1) gene leads to translation of PABPN1 diseased proteins. The normal PABPN1 gene contains 10 alanine encoding repeats, whereas the disease-causing genes are mutated to have 11 to 18 repeats in their sequence [20]. These deleterious mutations will translate into long alanine repeats in the N-terminus of the protein which are thought to be the cause of misfolded toxic proteins that lead to nuclear aggregates and cell death [25, 26].

As OPMD mainly affects the function of specific muscles, one would want to explore the function of this protein within muscle tissue. However, no muscle-specific function for PABPN1 has been found up to this date. The best-known function of the PABPN1 protein is the post-transcriptional modifications of the poly-A tail of the RNA transcript through interaction with poly(A) polymerase (PAP) enzyme to control polyadenylation of the mRNA, which determines its stability and its translation level [27]. Another suggested function for PABPN1 is transporting the RNA from the nucleus into the cytoplasm. Also, since PABPN1 is known to have a crucial role in regulating gene expression through alternative polyadenylation, and alternative polyadenylation was shown to affect muscle stem cells function, it can be hypothesized that non-functional mutants of PABPN1 which fail to regulate global gene expression within muscle cells, can affect the functionality of these cells by masking a certain polyadenylation site and leaving the alternative site available [28-31].

1.3 Current treatment options for OPMD

Currently, no effective treatments are available for OPMD patients [20]. It is commonly the symptoms of the disease and not the underlying cause that are subject to treatment. Progression of the disease can be lethal in late stages, but normally has little effect on the overall life expectancy of the patients [21, 32]. However, in more severe cases and, at more advanced stages of the disease, dysphagia, regurgitation, aspiration pneumonia and cognitive decline are prominent prior to death [33].

Nuclear aggregates form beta sheets, which contain polyalanine oligomers, and are very stable and resistant to enzymatic degradation [34]. These aggregates, which are caused by expanded polyalanine mutants, are therefore thought to be the main reason for cell death [35]. Protein aggregates in OPMD have been targeted and successfully reduced in mouse models, using different treatments such as cystamine [36], doxycycline, trehalose and guanabenz acetate (GA) [37], whereas in a *Drosophila* model of OPMD, intrabodies, which are antibodies that were designed to be expressed inside cells to target an antigen, have been shown to be effective [25].

Cell therapy is another treatment option that has been shown to be somewhat effective in OPMD patients. In a phase 1/2 clinical trial (ClinicalTrials.gov NCT00773227), autologous healthy myoblast cells were injected into the pharyngeal muscles of the OPMD patients. The results showed a cell-dose dependent improvement in swallowing, and a higher overall quality of life was observed in all 12 participating patients in this study [38].

1.4 Ribozymes

In 1982, Thomas R. Cech's research lab discovered the catalytic properties of RNA for the first time [39]. This was the beginning of the idea that an RNA molecule that can excise itself or catalyze cleavage reactions on other nucleic acid sequences can potentially be used as a therapeutic instrument. Ribozymes are small catalytic molecules that carry out enzyme-like activities in cells. Cis-acting ribozymes self-cleave (Figure 1), whereas trans-acting ones cleave other molecules and remain unchanged themselves (Figure 3) [40]. Unlike DNA, which is double stranded, RNA is a single polynucleotide chain that can fold back on itself and form tertiary structures. The ribose sugar in RNA (DNA has a deoxyribose sugar) allows the hydroxyl group in the second position to be deprotonated in basic conditions. This leaves the oxygen with a negative charge, which will attack the phosphate in the third position of the ribose sugar, and break the bond between the phosphodiester bond between the two RNA nucleotides, which cleaves the RNA at this site (Figure 2) [41].

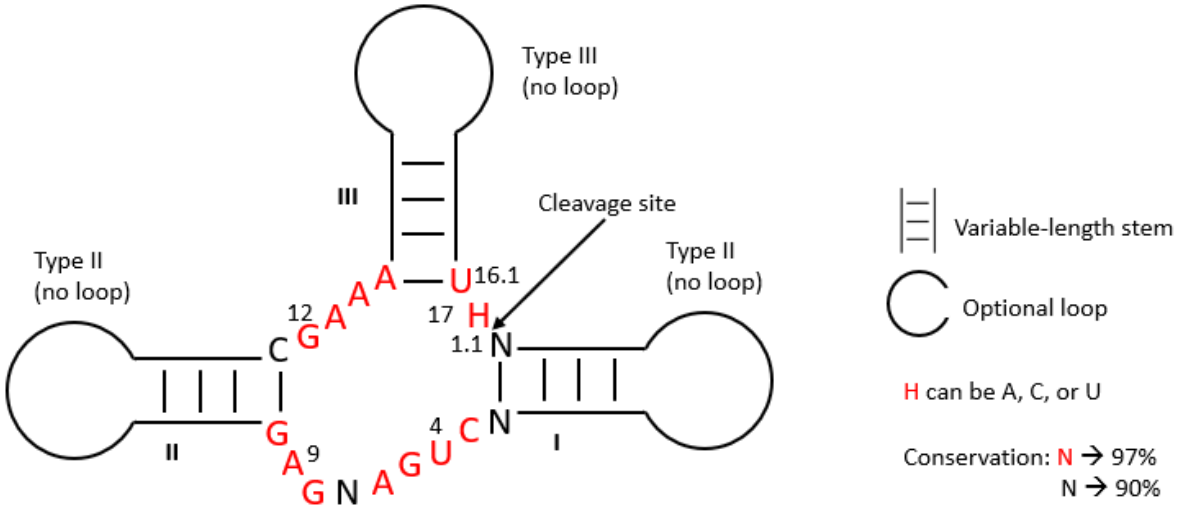


Figure 1. Self-cleaving hammerhead ribozyme consensus sequence

The optional loop determines the type of the hammerhead ribozyme numerically (type I, II or III). The red nucleotides show the most (at least 97%) conserved nucleotides in the catalytic core. N can be any nucleotide. The nucleotides in the loop are numbered clockwise and the nucleotide numbered 1.1 denotes the nucleotide that is the first one in the loop and helix I. The cleavage site is between helix III and helix I. Figure taken from [42].

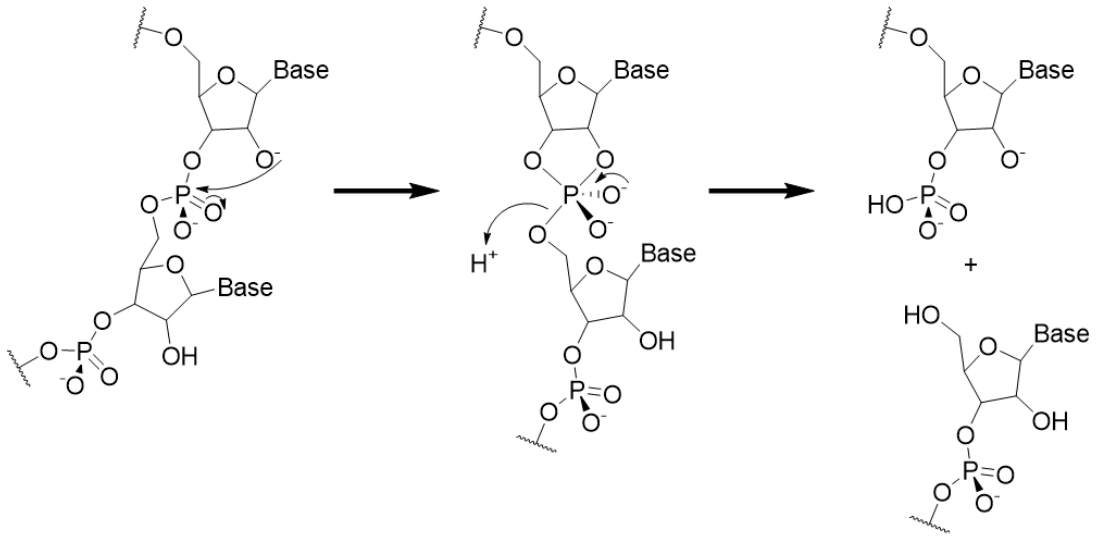


Figure 2. The reaction that takes place at the ribozyme cleavage site

RNA cleavage starts when the 2' oxygen attacks the phosphorus and causes the double bond with the oxygen to break (shown in the first image on the left). The bond between the two nucleotides of the RNA backbone is then broken when the bond between the phosphorus and the oxygen of the hydroxyl group attacks a proton and the cleavage is completed. At this stage, if the ribozyme is self-cleaving the ribozyme cleaves itself, while in trans-acting ribozymes the cleavage, of the mRNA that is bound to the binding arms of the ribozyme, is catalyzed.

1.5 RNA therapeutics for OPMD

OPMD is caused by mutations in a single gene PABPN1, which makes it a good target for gene therapy, as opposed to polygenic diseases, which may be more difficult to target. A 2017 paper by Malerba et al. shows that small hairpin RNA (shRNA) can be used to target the mutant PABPN1 mRNA and degrade it [19]. Simultaneous overexpression of codon-optimized wild type PABPN1, along with elimination of the diseased mRNA, was shown to alleviate OPMD symptoms and reduce the level of insoluble protein aggregates in cells, restoring normal muscle function and strength in a mouse model of OPMD and in OPMD patient cells [19].

RNA replacement therapy for OPMD using hammer head ribozymes or microRNAs (miRNAs) have shown promising results by targeting and reducing the level of the mutant mRNA in vivo in HEK293T, C2C12 OPMD model cells, as well as in *Caenorhabditis elegance* [43].

1.6 Inducible trans-cleaving hammerhead ribozymes

Hammerhead ribozymes (hhRzs) are a small family of self-cleaving ribozymes first discovered in viral plant pathogens. These ribozymes catalyze the excision of a section of themselves by breaking the backbone's phosphodiester bond [44]. The identification of the trans-cleaving

potential of these molecules has led to several studies with the aim of targeting different gene transcripts *in vitro* and *in vivo* [45]. A trans-cleaving hhRz is mainly comprised of target-binding arms and a catalytic core. The ribozyme stays inactive until the target strand that has a sequence complementary to the binding arm(s) of the hhRz is detected. Binding to the substrate allows the catalytic core of the ribozyme to catalyze cleavage of the target strand, which in turn will cleave the substrate next to the sequence NUC, NUA or NUU [46]. When designing a ribozyme, the oligonucleotide binding site (OBS) (where the ribozyme binds the substrate) is designed to interfere with the active conformation of the ribozyme. Binding of the substrate to the OBS lifts this repression and the ribozyme can cleave the substrate (Figure 3). Following this logic, it has been shown that these hhRzs can be designed to attack mutant transcripts of different length both *in vivo* and *in vitro* [47, 48].

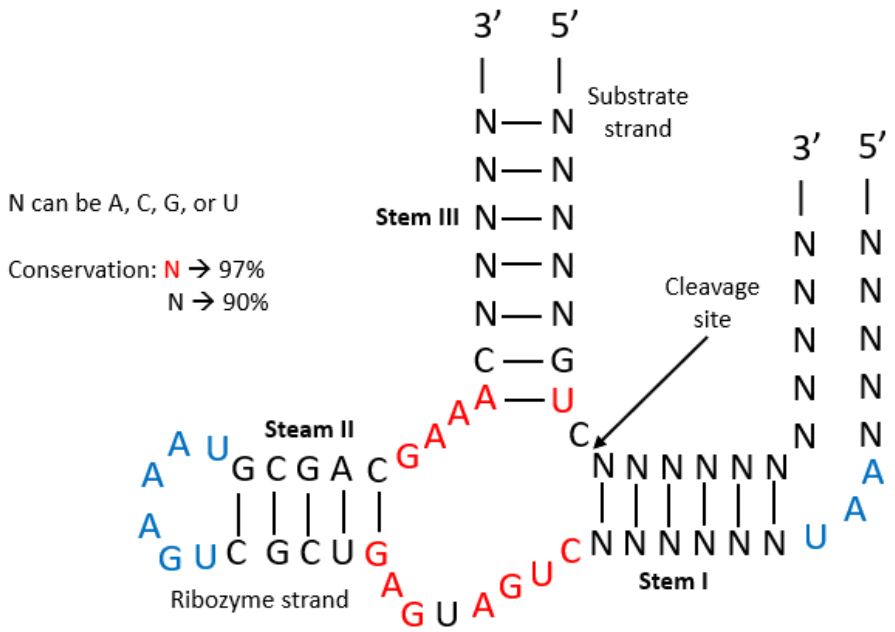


Figure 3. Sample structure of an inducible trans-cleaving hammerhead ribozyme

N can be any of the nucleotides. The substrate strand is bound to stem III and stem I of the ribozyme. The interaction of the blue nucleotides makes the formation of the tertiary structure of the ribozyme possible (not shown in this figure). The red nucleotides are highly conserved and form the catalytic core that makes the cleavage of the substrate mRNA molecule possible. The bound substrate is cut at the cleavage site. Figure taken from [42].

1.7 Objectives and hypothesis

The possibility of designing trans-cleaving hhRzs that can target RNA molecules of various length means that many mutant disease-causing mRNA can be targeted using customized ribozymes. As previously mentioned, trinucleotide expansion disorders that are associated with mutations within a single gene can be the best targets for RNA therapeutics. The problem of ribozymes targeting the normal transcripts can be solved if these ribozymes are designed to distinguish between the normal repeats (wild type) and the extended version (mutant transcripts).

Here we hypothesize that customized ribozymes designed to target PABN1 mutant alleles can successfully reduce the level of mutant mRNA and subsequently mutant protein in human cells (human embryonic kidney 293 cells).

The objectives of this research project are:

1. Test twenty-nine ribozymes selected by *Tri-cleaver* algorithm [48] against the wild type and mutant alleles of PABPN1 gene. Finding one hit out of 29 would be enough as it shows the potential of the algorithm to produce selective ribozymes and would provide valuable information for later optimizations.
2. Find the best ribozymes that will *selectively* cleave the two disease-causing mutant transcripts of PABPN1 gene with 13 and 17 alanine repeats (with minimal effects on WT with 10 alanine repeats) by optimizing the sequence of the best hits or combining effective ribozymes.

Chapter 2: Materials and Methods

2.1 Using a designed library of ribozymes to target two mutant PABPN1 transcripts

An evolutionary algorithm (EA), *Tri-cleaver* [48], was used in this thesis to design a library of trans-cleaving ribozymes targeting two mutant mRNA of PABPN1 gene, with 13 and 17 alanine repeats. The algorithm can generate RNA sequences that can potentially target the mutant transcript of many trinucleotide repeat expansion disorders.

The catalytic strand of the ribozyme binds to the target mRNA which is the substrate and cleaves it. The mRNA strands contain regions with common sequences which serve as ribozyme binding sites (RiBS) for the catalytic strands of the trans-cleaving ribozymes. The catalytic strands of the ribozymes contain oligonucleotide binding sites (OBS) that are reverse complementary to regions on the substrates. The OBSs prevent the ribozymes from forming their active conformations when they are not bound to the substrate.

In this project the ribozyme designs generated by the EA have GCG repeats in their sequence (part of the non-highlighted nucleotides in figure 4 and table 1) which are used as the actual OBS that will bind the target mRNA (Figure 4 and Figure 5B). The OBS interferes with the formation of the active form of the ribozyme, unless it is completely bound to the correct number of target repeats (Figures 5 and 6). The mRNA substrates in this project were two mutants of PABPN1 gene which contained 13 and 17 repeats of GCG, whereas the wild type only contained 10 alanine residues. Upon binding of the OBS region of the designed ribozymes to the mRNA with the target number of repeats, the ribozyme adopts its active conformation and cleaves the mutant mRNA.

DNA sequence of the ribozyme:

CgAAACCGGGCACTACAAAAACCAACTTTGGTGCGGCTGATGAGCGTCTGGTCAT
TAGTCGCTGCTGAAGTCGCTGCCTGCGTCGTCGGTCTGCCGGGCGCGAAAGCGGG
CGgtac

Transcribed to RNA:

~~CGAAACCGGGCACUACAAAAACCAAC~~~~UUU~~GGUGCGGCUGAUGAGCGUCUGGUC
AUUAGUCGCUGCUGAAGUCGCUGCCUGCGUCGUCGGUCUGCCGGGCGCGAAAG
CGGGCG~~guac~~

Figure 4. Looking at the DNA and RNA sequences of a ribozyme (Rbz8)

On the DNA sequence within the ribozyme plasmid the nucleotides highlighted in yellow are parts of the tRNA-Val promoter which is used for high expression levels of ribozymes *in vivo*. Highlighted in blue is the linker sequence. In green restriction cut site. In purple the binding arms of the ribozyme that will bind the target mRNA.

Table 1. Ribozyme sequences generated by *Tri-cleaver* algorithm [48].

(cg) in small letters and the sequence AAACCGGGCACTACAAAAACCAAC, highlighted in yellow, are parts of the tRNA-Val promoter which is used for high expression levels of ribozymes *in vivo*. These sequences are followed by a TTT linker sequence, highlighted in cyan. The ribozyme's binding arms sequences (which bind to the RiBS) are highlighted in purple, and the sequence of the catalytic core of the ribozymes is highlighted in gray. gtac sequence, highlighted in green, at the end of each strand is part of the KpnI restriction enzyme recognition and cut site (Ggtac^c).

Ribozyme ID	Ribozyme Sequences (5' to 3')
Rbz	cgAAACCGGGCACTACAAAAACCAACTTTGTTTTCGCTGATGAGTCCGGACC CTTTGCCCTTGTCTGCGTCGTCCTAGTCCCATTCCGGGCGAAGGTCCGgtac
Rbz0	cgAAACCGGGCACTACAAAAACCAACTTTGCTCGTCCTGATGAGGGGTGGC GCAGACGCCGCCGCCGAAGCCGTTGTCAATGTCGCCACCCTGACCGCCTCG AAAAGGCTAGgtac
Rbz1	cgAAACCGGGCACTACAAAAACCAACTTTTAGGGTCCTGATGAGCGGCACT TCAAGTGCCGCCGCCGAAGTCGCCGCTAATGCTGCCCTGTTGGGTGTTGCG AAAGGGCGGGgtac
Rbz2	cgAAACCGGGCACTACAAAAACCAACTTTGGCGCGGCTGATGAGTGCTGCC GGCATAGCCGCCGTCGAAGTCGCCGCTAATGCCGCTCGTGTGAGCGGCGCG AAAGTGGGCGgtac
Rbz3	cgAAACCGGGCACTACAAAAACCAACTTTGGCGCGGCTGATGAGTTCTTTC AAGTCAGCCGCCGTCGAAGTCGCCGCCAATGCCGCCCGGGGGGGAGGGGC GAAAGCGGGCGgtac
Rbz4	cgAAACCGGGCACTACAAAAACCAACTTTTGGCTTGCTGATGAGGTGGGCG CGCGGAGCCGCCGTCGAAGCCGTTGTCAGCGCCGCCCCCCTGGGTCCATCG AAAGTCTCGGgtac
Rbz5	cgAAACCGGGCACTACAAAAACCAACTTTGAGTAGGCTGATGAGGATGGTG GGAGCCGCCGCCGTCGAAGCCGTTGTCAACGCCGTCCCGTCCGGCCATTCG AAACTCGTTGgtac
Rbz6	cgAAACCGGGCACTACAAAAACCAACTTTGTCTCTGCTGATGAGGGTGCCG CTAGACGCCGTCGCCGAAGTCGCCGCTAATGCTGCCCCGAGATGGTGCTCG AAAGGCGCGGgtac
Rbz7	cgAAACCGGGCACTACAAAAACCAACTTTGCTGTTGCTGATGAGGGGTGGC GCAGACGCCGCCGCCGAAGCCGTTGTCAATGTCGCCACCCCGGCCGCCTCG AAAGTTGGTGgtac
Rbz8	cgAAACCGGGCACTACAAAAACCAACTTTGGTGCGGCTGATGAGCGTCTGG TCATTAGTCGCTGCTGAAGTCGCTGCCTGCGTCGTCGGTCTGCCGGGCGCG AAAGCGGGCGgtac
Rbz9	cgAAACCGGGCACTACAAAAACCAACTTTCGGCTTGCTGATGAGGGGGGCG CGCGGAGCCGCCGTCGAAGCCGTTGTCAGCGCCGCCCTCTTTTGTCCCTCG AAAGTCTCGGgtac

Rbz10	cgAAACCGGGCACTACAAAAACCAACTTTGCCTCTGCTGATGAGGCGCCCG TTTGACGCCGTCGCCGAAGTCGCCGCTAATGCTGCCCTCAAAAGGGTGTCG AAAGGCGCGGgtac
Rbz11	cgAAACCGGGCACTACAAAAACCAACTTTGGGTCCTCTGATGAGTCAGTGA CCGAACGCCGCCGCCGAAGCTGTCGCTAACGTCGCCAGTTATCCACTGGCG AAACCGTCCGgtac
Rbz12	cgAAACCGGGCACTACAAAAACCAACTTTGGTGCGGCTGATGAGGACGGCG AGCGGTGCCGTCGTCGAAGCTGCTGTCAGTGCTGTCCTCAACAGCCGTCCG AAAGCGGGCGgtac
Rbz13	cgAAACCGGGCACTACAAAAACCAACTTTGCCAACGCTGATGAGGGGTGGC CAAATCGCCGCCGCCGAAGCCGTTGTCAATGTCGCCACCCCGGCCGCCTCG AAAGGTGGAGgtac
Rbz14	cgAAACCGGGCACTACAAAAACCAACTTTGCTTTGGCTGATGAGTTATCTCA CAACAGCCGTCGTCGAAGTCGCTGCCAATGTCGTCCGGAAAGGGATGGCG AAAGTTTCCGgtac
Rbz15	cgAAACCGGGCACTACAAAAACCAACTTTGTTCTCGCTGATGAGCTATCTCG TGGTCGCCGCTGTCGAGGCCGCTGTCCACGTCGTTGCCCAGTAGGTAGCGA AAGTCTTCGgtac
Rbz16	cgAAACCGGGCACTACAAAAACCAACTTTGGCGTGGCTGATGAGCTGACTC GATACAGTCGTCGCCGGAGCCGCTGTCTATGTCGTCAGTCTCCAGTCGGCG AAAGCGGGCGgtac
Rbz17	cgAAACCGGGCACTACAAAAACCAACTTTGGCGCGGCTGATGAGCGCCAAG ATGCGTGTCGTCGCTGAAGCCGCCGCTTGCGTCGTCAGCCGGCTTGGTGCG AAAGTGGGCGgtac
Rbz18	cgAAACCGGGCACTACAAAAACCAACTTTGGCGTGGCTGATGAGGGCGGCG GTACTAGCCGCTGTCGAAGCCGTTGTCAATGCCGTCCTCGTGGGCCGCTCG AAAGCGGGTGgtac
Rbz19	cgAAACCGGGCACTACAAAAACCAACTTTGGCGCGGCTGATGAGTGGCCTA GTCCAAGCCGCTGTCGAAGTTGCTGCTAATGCTGCCACTCTGTAGGCTGCG AAAGTGGGCGgtac
Rbz20	cgAAACCGGGCACTACAAAAACCAACTTTGGCGTGGCTGATGAGCGGCCTT TAATCGGCCGTCGCCGAAGTCGCCGCCCATGCCGCCTACAGTGGGGCTGCG AAAGCGGGCGgtac
Rbz2840	cgAAACCGGGCACTACAAAAACCAACTTTCACCCTCCTGATGAGCTGGCGG CCATAGGCCGCCGCCGAAGTCGTCGCCAATGCCGCCAGTTAGTCGCCGGCG AAACCGGTCGgtac
Rbz2841	cgAAACCGGGCACTACAAAAACCAACTTTCATTCTCCTGATGAGGGTGCCA CCCTCCGCCGTTGTCGAGGTTGCCGTTCGCGCCGTCTCTGGACGGCGCCCG AAACCGGTTGgtac
Rbz2842	cgAAACCGGGCACTACAAAAACCAACTTTCACCTTCCTGATGAGGGGCTGC GGGCTCGCCGTCGTCGAGGTTGCTGTCTATGTTGTCAGGACGGTGGCTCCG AAACTAGTCGgtac
Rbz2843	cgAAACCGGGCACTACAAAAACCAACTTTCATCTTCCTGATGAGACTGGCG GAATGCGCCGCCGCTGAGGTCGTTGTCCATGCCGTCAGACCCCGTCGGTCG AAACTGGTCGgtac

Rbz2844	cgAAACCGGGCACTACAAAAACCAACTTTCGCTCTCCTGATGAGAGCTGTTTGACCAGTCGCTGTCGGAGTCGCTGCCTACGCCGTTCGTCGGGATAGCTCGAAACCAGTCGgtac
Rbz2845	cgAAACCGGGCACTACAAAAACCAACTTTCATCCTCCTGATGAGCGGTTCTTTCAATGCCGCCGTCGGGGTTGCCGTCTACGCTGTCCAAATATGGACTGCGAAACCGGTCGgtac
Rbz2846	cgAAACCGGGCACTACAAAAACCAACTTTCACCCTCCTGATGAGTTGCGCTGGTTTTGTCGCTGTTGGAGTTGCTGCCAATGCCGCCCTTGTCGGCGCGGCGAAACCGGTCGgtac
Rbz2847	cgAAACCGGGCACTACAAAAACCAACTTTCACCCTCCTGATGAGCTGCTCAGCGCTTGCCGCCGTTGAGGTTGCCGTCAATGCTGCCTTACCTTGGGCGGCGAAACCAGTCGgtac

2.1.1 Mechanism of action of hammerhead ribozymes in cleaving mutant mRNA

The algorithm designs the catalytic strand of the ribozymes to bind downstream of the GCG repeats on the mRNA of both wild type and mutant type PABPN1 genes. The catalytic strands are designed to be active only when the entire OBS is bound to the trinucleotide repeats on the mRNA. The length of the OBS sequence on the catalytic strand of the ribozyme, which is reverse complementary to the repeat sequences on the mRNA, is longer than the length of the repeats on the wild type gene, but shorter than or equal to the length of the mutant sequences. Recall that for the ribozyme to fold into its active conformation both the RiBS and the OBS sequences must be bound. Therefore, the ribozyme is only active when it is bound to the mutant mRNA, because the entire OBS region, which is the same length as (or shorter than) the trinucleotide repeats, is bound. When the ribozyme binds to the wild type, the OBS region is only partially bound to the repeat region, and the ribozyme does not become active (Figures 5 and 6).

2.1.2 Selectivity and efficiency of ribozymes

The designed ribozymes are highly specific, meaning that the ribozyme binding site must be present on the mRNA. Also, the OBS must be reverse complementary to the GCG repeat sequences on the transcripts. The secondary structures of these trans-acting ribozymes are predicted using Vienna RNA folding package [49]. These ribozymes are expected to be selective and efficient in binding to the target mRNA and cleaving only the longer mutant repeats.

Figure 3 depicts the cleavage mechanism of a trans-cleaving ribozyme such as Rbz8. For a given sequence of a ribozyme, there can be many possible inactive folded structures, where the oligonucleotide binding site interferes with the formation of the active structure of the ribozyme. This means that the ribozyme can fold and form various stems and loops, but the catalytic core and the binding arms will not be available in any of these conformations, and the ribozyme remains inactive. A possible structure for Rbz8 sequence, which has the lowest free energy is shown in

figure 5A. In presence of the target mRNA, the long repeats of the mutant transcript will bind to the ribozyme's OBS and loop II will form. This will allow the formation of the ribozyme's catalytic core and release of the binding arms (Figure 5B). In contrast, the wild-type mRNA does not contain a long enough GCG sequence to cause the formation of loop II and putting the ribozyme in active conformation.

Upon binding of a mutant transcripts (13 or 17 alanine codon repeats) to the OBS and formation of the active conformation, the mRNA will bind the binding arms and the cleavage reaction will be catalyzed by the ribozyme, leading to the degradation of the mutant transcript (Figure 5B). The assumption that, only the binding of the expanded GCG sequence of the PABPN1 mutant mRNA (and not the wild type sequence containing 10 alanine codon repeats), would result in the formation of the active ribozyme conformation, was shown to be true to some extent for Rbz8 and Rbz5.

Ten alanine repeats are not enough to compete with the base pairing of the nucleotides and the formation of the two helices in the inactive form of the ribozyme, as can be seen in Figure 5. Therefore, in the absence of 13 or more GCG repeats, the ribozyme will fold on itself. This conformation will interfere with the formation of the second loop and the catalytic core; hence the ribozyme remains inactive (Figure 5A). Ten repeats are not enough to unfold the ribozyme from this conformation; however, more repeats will unfold the ribozyme and cause the formation of loop II and the catalytic core subsequently (Figure 5A and Figure 6).

It is important to note that the base pairing of RNA molecules does not always follow the Watson-Crick rules. The non-Watson-Crick base pairing, also called "wobble" base pairs, are essential to RNA molecule secondary structures and are present in almost all types of RNA, across all domains of life [50]. The most important of these wobble base pairs is the G.U base pairing, which occurs naturally and plays a crucial role in the formation of the secondary structure of different classes of RNA [51]. More specifically, the G.U wobble pair in the active site of many classes of ribozymes, has been shown to be responsible for the reactivity of the ribozymes; replacing it with a conventional Watson-Crick base pair disrupts the function of the ribozyme.

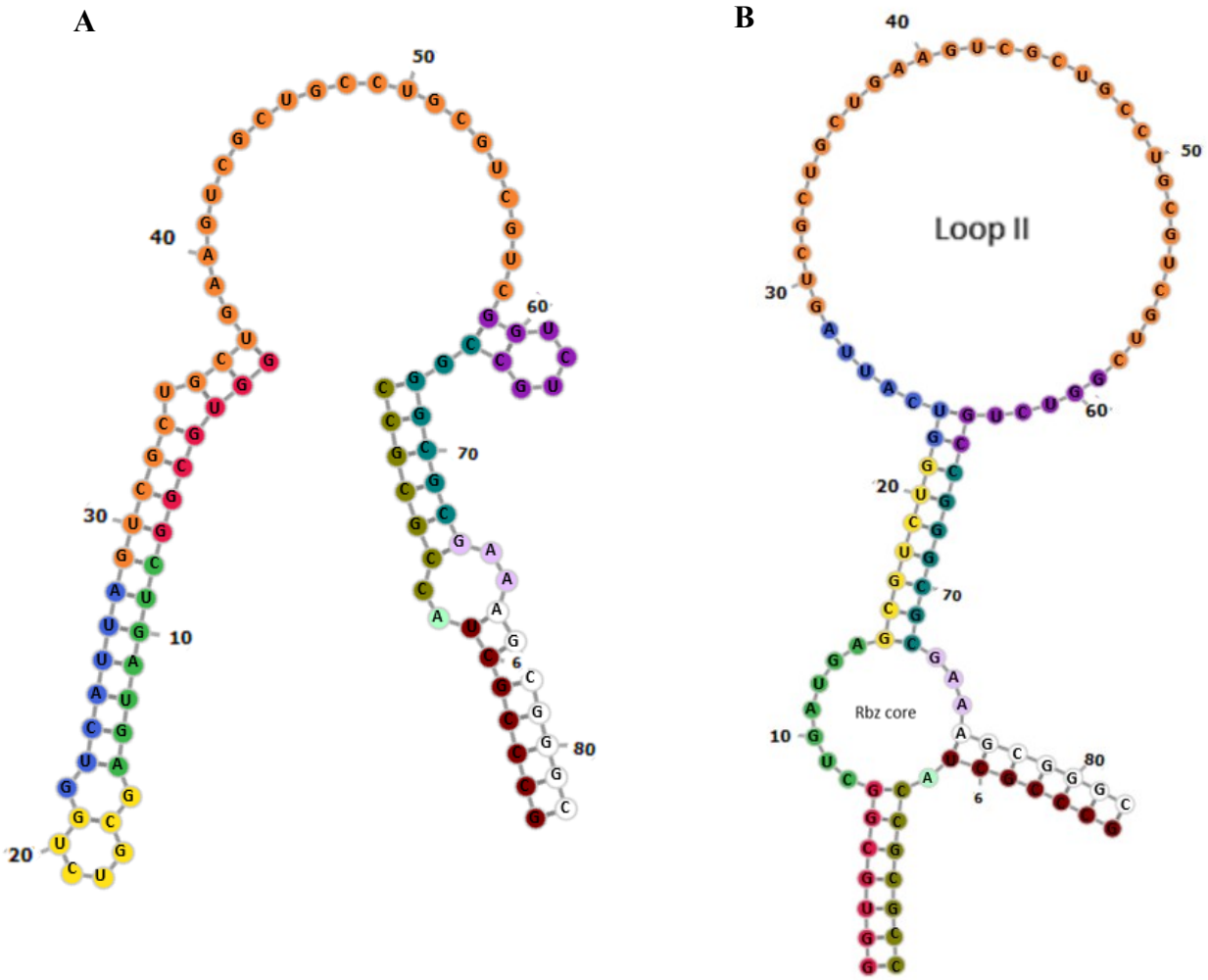


Figure 5. Inactive and active conformations of the Rbz8 sequence

In panel A, an inactive conformation of Rbz8 with free energy of -46.70 kcal/mol is depicted. This free energy is the lowest amongst the possible conformations generated using two RNA folding web servers, FORNA and UNAfold. As shown in A, binding of the WT PABPN1 mRNA with 10 GCG repeats to the OBS of the ribozymes (in orange, purple and blue, respectively) is not enough to form the loop (loop II) and therefore the stems of a hammerhead trans-cleaving ribozyme and the catalytic core which allows cleavage do not form. Hence, the inactive ribozyme will not be able to cleave the WT mRNA. However, if the number of GCG repeats on the mRNA that bind the OBS is larger (13 or 17) loop II, shown in panel B can form and the inactive ribozyme, with free energy of -37.59 kcal/mol, can unfold to form the active conformation (panel B). The active ribozyme with the loop, stems and the catalytic core will result in the cleavage of these target mRNAs (Image generated using FORNA web server [52]).

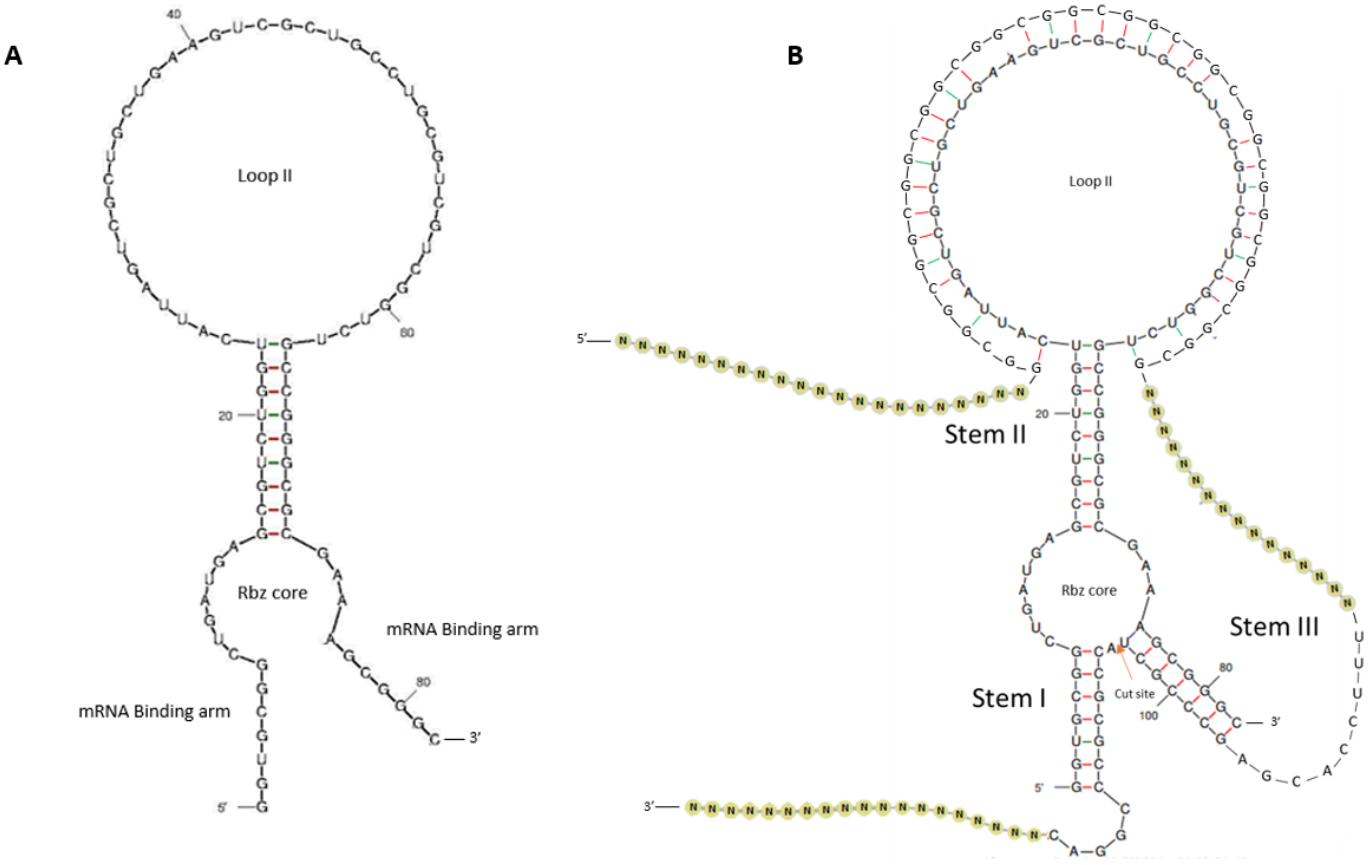


Figure 6. Active conformation of Rbz8 with target mRNA

The active form of Rbz8 is shown in A where the three stems (I, II, and III), the catalytic core and the binding arms form (the bound target is not shown here). In B, the N in the target sequence denotes any of the nucleotides. Binding of the target, which contains more than the 10 alanine (GCG) repeats in wild-type PABPN1 sequence, to the ribozyme's OBS, will result in the active structure of the ribozyme that will allow cleavage of the target at the cut site. Mismatches between the OBS and the target sequence are intentional and will prevent long stretch of double stranded RNA which can be targeted for cleavage by enzymes such as Dicer. The black bonds between the nucleotides are phosphodiester bonds that make up the backbone of the ribozyme and the mRNA sequence. The red hydrogen bonds show the base pairing between Guanine and Cytosine (Watson-Crick base pairing) while the green hydrogen bonds are non-Watson-Crick G.U wobble base pairs (Image generated using UNAFold web server, 2022).

2.2 Cloning the ribozyme sequences (double stranded DNA) into plasmids

The coding sequences of ribozymes were generated using the *TriCleaver* evolutionary algorithm, and
synthetically generated ribozymes were cloned using the expression vector pUC-KE-tRNA-CTE (derived from pUC 19) with ampicillin resistance suitable for mammalian cells obtained from Nawrot, Barabara [53]. pUC-KE-tRNA-CTE contained $tRNA^{Val}$ promoter which drives the expression of a CTE helicase associated ribozyme encoded on the plasmid. The DNA encoding the ribozyme was cut out of the plasmid and replaced by the ribozymes used in this project.

Initially, 29 ribozymes were designed to target the PABPN1 gene. Top and bottom strands of every ribozyme were generated by Integrated DNA Technologies (IDT). The two strands were then annealed using the supplier's protocol. Each double stranded ribozyme sequence and the expression vector were digested with KpnI-HF (NEB) and BstBI (NEB) using the supplier's protocols. Every DNA sequence encoding a ribozyme was subsequently ligated into the expression vector (pUC 19) by using the Quick Ligase kit (NEB) and the corresponding protocol. The sequences of the ribozymes containing plasmids were then verified using Sanger sequencing and the PCR primer sequence (5′-CGCCAGGGTTTTCCCAGTCACGAC-3′).

The generated plasmids were then used to transform bacterial cells Agilent Technologies XL 10-Gold Ultracompetent cells following the supplier's protocol of transformation. The resulting colonies were picked and cultured for subsequent miniprepping.
The plasmids were then extracted using QIAprep Spin Miniprep Kit (QIAGEN) using the corresponding protocol but eluted in lower volumes of 20 µl, instead of the recommended volume of 50 µl, to obtain higher concentration for sequencing. The miniprepped plasmids were sent to McGill Genome Center for sequencing to confirm the presence of the correct ribozyme sequence in each plasmid.

2.3 Transfecting HEK293E cells with PABPN1-gene-carrying and ribozyme-carrying plasmids

The wild type gene of PABPN1 carries 10 alanine coding GCG sequences while the two disease-causing mutants have 13 and 17 GCG repeats respectively. The plasmids were kindly provided by Rouleau lab [43, 54]. Plasmids were prepared by cloning cDNAs of PABPN1 wild type and mutant gene into pEGFP-C2 vector (Clontech, Palo Alto, CA, USA) as described in Messaed *et al*. This resulted in each plasmid coding a PABPN1-GFP fusion from which the fluorescent signal can be used to confirm transfection.

To test each of the ribozymes, Human Embryonic Kidney Cells (HEK293E) were cultured in DMEM (Invitrogen) containing 10% fetal bovine serum in cell culture incubator at 37 C. The cells were seeded in 12-well plates and were transfected at 70% to 80% confluency using the jetPRIME transfection reagent (Polyplus) and the corresponding supplier's protocol. The patterns based on which the experiments were set up and the amount of plasmid DNA used for transfection is shown in the following table (Table 2).

Table 2. General setup for transfecting wells in a 12-well plate.

2 µg Wild type PABPN1 plasmid DNA	1 µg Wild type PABPN1 + 1 µg Ribozyme plasmid DNA	1 µg Wild type PABPN1 + 1 µg control Ribozyme plasmid DNA	1 µg Ribozyme plasmid DNA
2 µg Mutant 13 GCG PABPN1 plasmid DNA	1 µg Mutant 13-Ala PABPN1 + 1µg Ribozyme plasmid DNA	1 µg Mutant 13 GCG PABPN1 + 1 µg control Ribozyme plasmid DNA	1 µg Ribozyme plasmid DNA
2 µg Mutant 17 GCG PABPN1 plasmid DNA	1 µg Mutant 17-Ala PABPN1 1ug + Ribozyme 1ug plasmid DNA	1 µg Mutant 17 GCG PABPN1 + 1 µg control Ribozyme plasmid DNA	Non-transfected cells

2.4 Fluorescent imaging of transfected cells

Images of each well on the plate were taken using Invitrogen EVOS fluorescence microscope. The presence of fluorescent signals confirms the efficiency of transfection in each well, since the expression vectors of wild type gene and two mutants all contain GFP markers. At least three different fields were visualized, and images were captured from these fields representing the overall transfection efficiency.

2.5 RNA extraction

All RNA extraction steps were performed in nuclease-free environments inside a biosafety cabinet using RNaseZap (Invitrogen). 300 µl of TRIZOL (Life Technologies) was used to extract the cells from each well of 12 well plates at different time points (24hrs, 48hrs, and 72hrs post-transfection). 70 µl of chloroform was added to promote phase separation to separate the RNA from DNA and proteins. The mixture was then vigorously vortexed for 15 seconds and incubated at room temperature for 2 to 3 minutes before being centrifuged for 15 minutes at 12000 rpm. The aqueous phase which contained the RNA was then collected and 150 µl of isopropyl alcohol was added to precipitate the RNA. The samples are then incubated at room temperature for 10 minutes. Alternatively, RNA can be precipitated overnight in -20 degrees Celsius for larger yield.

After incubation, the samples are centrifuged for 10 minutes at 12000 rpm and the supernatant is discarded. 300 µl of 75% ethanol was added and the samples were incubated in -80 degrees Celsius overnight. The samples were then centrifuged at 12000 rpm for 5 minutes, ethanol was discarded and the ethanol wash step was repeated. After discarding the ethanol in the second step the caps were left open for about 10 minutes for the residual ethanol to evaporate. The RNA pellet was then re-suspended in 18ul of nuclease-free water and the samples were stored in -80 degrees Celsius.

2.6 Protein Extraction

The cells were collected from each well of the 12 well plates at different time points (24hrs, 48hrs, and 72hrs post-transfection) by scraping, and were centrifuged at 6000 rpm for 5 minutes. The supernatant was discarded, and the cells were washed by adding 300 µl PBS and centrifuging at 6000 rpm for 5 minutes. 60 µl sodium dodecyl sulfate utilizing buffer (SUB) containing 8M urea, 2% β-mercaptoethanol, and 0.5% SDS was then added to each sample. The samples were sonicated and the concentration of each sample was measured using the standard Bradford Assay.

2.7 Quantitative RT-PCR analysis

cDNA was synthesized from the purified RNA samples using SuperScript Vilo (Thermo Fisher Scientific) and mixed with the PABPN1 probe with the sequence 5′-TCGAGGGTGACCCGGGGGA-3′ (TaqMan Gene Expression Assays PABPN1, Applied Biosystems Applied Biosystems, Hs01091143-g1) according to the supplier's protocol. Relative gene expression was calculated by normalizing expression against the reference endogenous gene. The endogenous control (the reference gene) for all experiments was RNA polymerase II probe (Applied Biosystems, TaqMan Gene Expression Assay probe against human POLR2A- 4331182 Hs00172187_m1) and all experiments were done in triplicates. qPCR experiments were conducted using 96 well-plates in QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) using TaqMan reagents to detect the targetsequences. All experiments were performed at 50°C for 2 min, 95°C for 2 min, and then 40 cycles of 95°C for 1s and 60°C for 20s. The threshold crossing value was noted for each transcript and normalized to the internal control with the internal control RNA polymerase II (GGGGCGGCCTCCCTCAGTCGTCTCTGGGT ATTTGATGCCACCCTCCGTCACAGACATTCGC).

Relative quantification method (Comparative CT method (ΔΔCT)) was used for relative quantification of mRNA in each sample based on Bio-rad real-time PCR application guide. Subsequent data analysis and comparison were performed using QuantStudio Real-time PCR software (Applied Biosystems). In this method the value of CT is the threshold at which fluorescence signal from enough amplicons is detected. Hence a lower CT value shows the presence of higher concentration of starting material (the target cDNA). ΔΔCT method assumes the efficiency of amplification of the reference gene and the gene of interest to be at a hundred percent. The expression of both the target gene and the endogenous control are measured in all control samples and all target samples. This provides us with four CT values (target gene in target samples, target gene in control samples, control gene in target samples and control gene in control samples). The CT of the target gene is normalized to the CT of the reference gene for both the target samples and the control samples (ΔCT = CT (target) – CT (reference)). This provides us with two ΔCT values. To obtain the ΔΔCT value the ΔCT of the samples is normalized to those of the corresponding reference gene by simply subtracting the ΔCT of the reference samples from the ΔCT of the target gene samples. The ΔΔCT value shows the change in the expression of the gene of interest normalized for any differences in amount of cDNA in the input material. We then use $2^{-\Delta\Delta CT}$ to calculate the relative quantification (RQ) which is the fold difference in expression

level of the target gene compared to the endogenous gene. RQ values were plotted against the sample to represent the relative gene expression [55].

To plot the relative gene expression in figure 16 (Appendix), the RQ values for each transcript type (WT, M13, or M17) in presence of a ribozyme, is obtained by setting the control expression value to 1. For example, to determine the transcript level of wild type PABPN1 in presence of Rbz8, the transcript level of wild type with the control ribozyme is set to one (100%) and the RQ of the level of the wild-type transcript in presence of Rbz8 is recorded relative to this control. This means that the effect of each ribozyme on each of the mutant transcripts (M13 or M17) and the wild-type transcript (WT) is looked at independent of the other two. These values can then represent the relative expression level of that gene in presence of a certain ribozyme. These values are obtained from the QuantStudio™ Real-Time PCR software v1.3 and the subsequent calculations of their standard errors are included in figure 15 in the appendix.

2.8 Western blot analysis

The acrylamide gels were prepared using the TGX™ FastCast™ Acrylamide Kit, 10% (Bio-Rad). Equal concentrations of each protein sample were loaded on the gel. The proteins were transferred onto nitrocellulose membranes using the Trans-Blot Turbo Transfer System (Bio-Rad). The blots were incubated with PABPN1 antibodies (Abcam, ab75855) (1:2000) and milk (5%, w/v) overnight before being developed using the Clarity western Blotting Substrate (Bio-Rad) in the ChemiDoc System (Bio-Rad). Total protein content was visualized using the same system prior to staining the blots with chemiluminescence solutions. Total protein content was used in all experiments to normalize the protein content of each well. Although actine antibody was applied to the blots in the initial experiments and its bands are present on the blots, it was not used as a means of normalization of protein contents. All independent experiments were done in triplicates. Protein content of each band was quantified using ImageJ software [56].

To quantify the bands on each blot, the intensity of the pixels of each band in a defined area was measured. ImageJ software measures the maximum, minimum, and the mean intensity. The mean values were used for all calculations [56]. The intensities were measured over the same area (enclosed box) for bands on each blot. After measuring the pixel intensity of each band, the pixel intensity of the background was subtracted from the measured mean intensities. Next, the pixel intensity of each of the total protein content columns associated with the bands (same area for all columns on one plot) was measured and the pixel intensity of the background was subtracted from these values. The mean intensity of each band was then divided by the mean intensity of the corresponding total protein column to ensure the band intensities are normalized to the total protein loaded on the gel and are comparable.

Since each set of samples were run in duplicates, the values obtained from each duplicated blot were then normalized to a percentage. The average of the two percentage values (one from each blot) was calculated and used to generate the relative protein expression graphs. The sample calculations for one of the blots is presented in figure 11 in the appendix.

Chapter 3: Results

3.1 Fluorescent imaging of transfected cells

Twenty-nine ribozymes were screened to identify those that selectively cleave the mutant PABPN1 transcripts (M13, M17) with minimal effects on the levels of the PABPN1 wild type (WT) protein. Following the transfection of the HEK293E cells with two plasmids that code for PABPN1-GFP (wild-type of both mutants) and one ribozyme, fluorescence imaging of the cells was performed to confirm transfection by visualizing the fluorescence signal of the PABPN1-GFP fusions. The images were taken of all duplicates to ensure comparable transfection efficiency in plates that were used for either protein or RNA extraction (Figure 7). The signal was used only to visualize successful transfection of the cells and to eliminate the samples that contain aggregates of dead cells. The plasmids containing M17 and wildtype PABPN1 fused to GFP were different from the plasmid that contained M13, meaning M13 was being expressed under a different less efficient promoter than WT and M17. Therefore, the signal from M13 and hence the RNA and protein expression of M13 is less than the wildtype and M17, and this is reflected in blot and RT-qPCR results. The following images (Figure 7), which correspond to Ribozymes number 5 (Rbz5) and the control ribozyme (MJD3) samples, show the change in the GFP signal in presence (+) and absence (-) of the ribozyme. Imaging was done for samples of every ribozymes in every experiment.

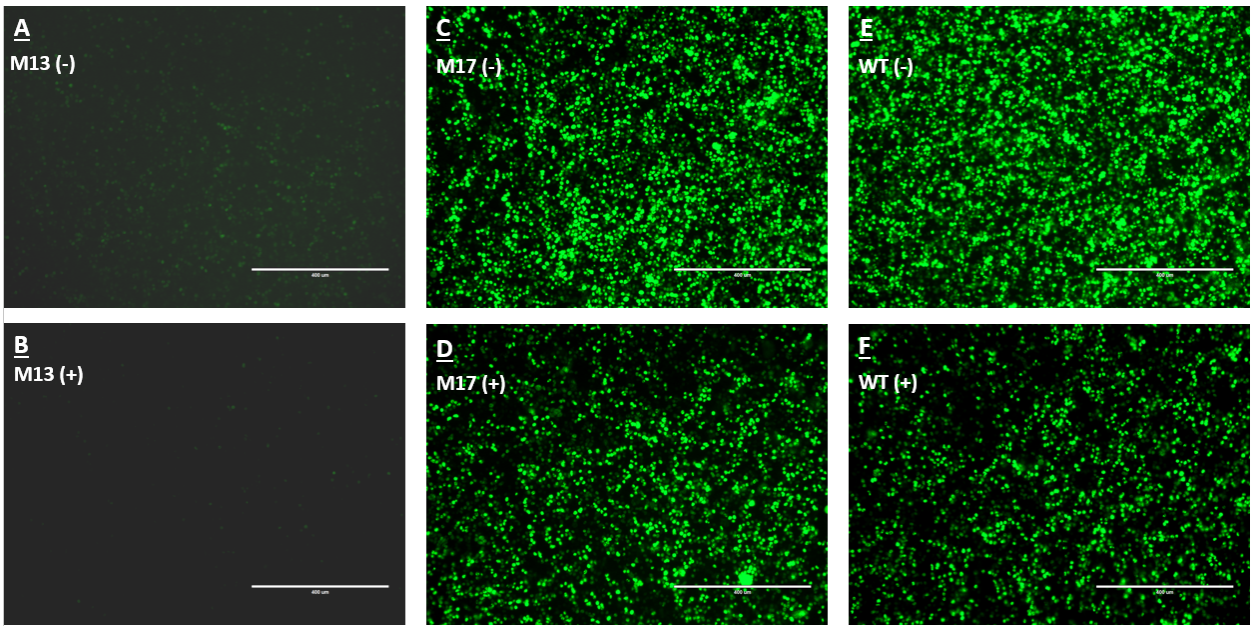


Figure 7. Fluorescent imaging of PABPN1 wild-type (WT), Mutant 13-Ala (M13), and Mutant 17-Ala (M17) 24hrs post-transfection

The signal from the GFP fusion protein is an indication of successful transfection. It can also be a rough estimate of the change in protein levels. Images A, C, E show M13, and M17, and WT protein expression in the absence of Rbz5 (-) respectively, and images B, D and F show the protein expression of the same set in presence of Rbz5 (+).

3.2 Protein expression levels in the presence of the mutant transcript-targeting ribozymes

After imaging each well and extracting the proteins, the samples were quantified. Protein gel electrophoresis was performed, and the blot results of all 29 ribozymes were generated using PABPN1 antibodies (Figure 10, Appendix). The bands on the blots are representative of the relative amount of PABPN1 protein present in cells at the time of extraction. Hence, these results show the effects of each tested ribozyme on protein levels of PABPN1 wild-type (W), mutant-13 Ala (M13), and mutant-17 Ala (M17). Blot results of samples of all tested ribozymes can be found in the appendix (Figure 10, Appendix). These results were used to categorize the ribozymes into three distinct groups based on their ability to knock down PABPN1 protein and their selectiveness:

1. Ribozymes that cleave *all* three transcripts, the PABPN1 wild type, mutant 13-Ala and mutant 17-Ala: non-selective, efficient.

2. Ribozymes that do *not* cleave any transcript: non-efficient, non-selective.

3. Ribozymes that cleave *mostly* the mutants (either mutant 13-Ala and/or mutant 17-Ala) and do *not* cleave the wild type: efficient and selective.

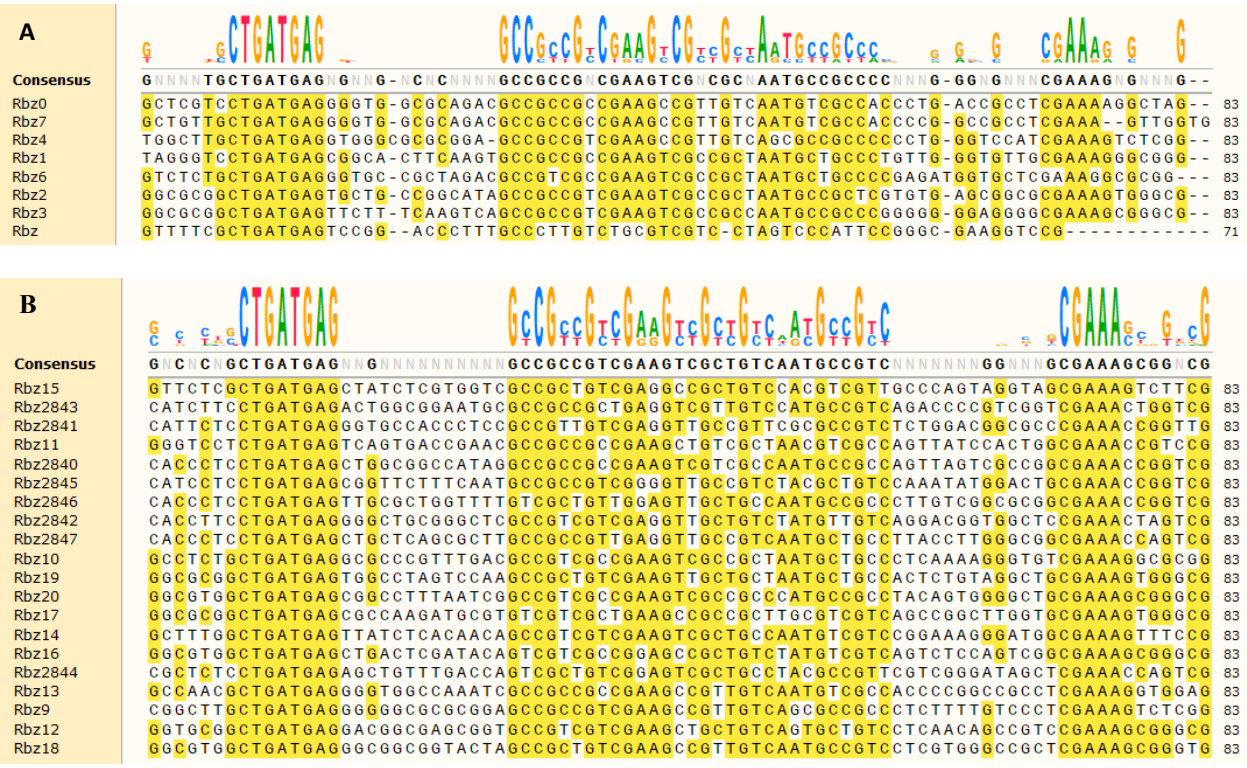
Table 3 shows all twenty-nine ribozymes organized into said groups. Ten ribozymes were found to reduce PABPN1 protein levels in cells, of which two (Rbz5 and Rbz8) were not only efficient, but also somewhat selective.

Figure 8 shows the alignment of the ribozymes grouped in table 3. Since the catalytic core of the hammerhead ribozyme is conserved, all ribozymes have this sequence in common (shown in table1 as well). There is a high degree of conservation in the trinucleotide binding region, which is the OBS of the ribozymes, because they must retain a degree of complementarity to the mRNA. The nucleotides in the binding arms of the ribozymes show a degree of similarity, since this is the part of the ribozyme that needs to be complimentary to the PABPN1 mRNA. However, there is a high degree of variability in the ribozymes' stems, both within and between the groups. The stems are essential for the catalytic core to from and cleave, but there is room for variations within these regions. For group C ribozymes, which are selective and efficient, the sequences appear to have more similarities in their stems (in addition to other conserved regions), but due to the small sample size, there is little basis to make confident general inferences about which sequence elements are most important.

Table 3. Tested ribozymes organized into three groups based on their efficiency and selectivity to knock down PABPN1 mutant proteins

Efficient non-selective ribozymes **A**	Non-efficient non-selective ribozymes **B**		Efficient and selective ribozymes **C**
• Rbz	• Rbz9	• Rbz19	• Rbz5
• Rbz0	• Rbz10	• Rbz20	• Rbz8
• Rbz1	• Rbz11	• Rbz2841	
• Rbz2	• Rbz12	• Rbz2840	
• Rbz3	• Rbz13	• Rbz2842	
• Rbz4	• Rbz14	• Rbz2843	
• Rbz6	• Rbz15	• Rbz2844	
• Rbz7	• Rbz16	• Rbz2845	
	• Rbz17	• Rbz2846	
	• Rbz18	• Rbz2847	

The efficient but non-selective ribozymes such as Rbz3 and Rbz6 can cleave all or at least two of the PABPN1 transcript variants (M13, and M17) very effectively, but they also consistently efficiently cleave the PABPN1 wild-type transcript (Figure 10 (C), (D) and Figure 12, Appendix). The non-efficient non-selective ribozymes do not reduce the protein levels of PABPN1 transcripts. Ribozymes 18, 19 and 2844, for example, do not have a knock down effect on any of the PABPN1 proteins, as can be seen in the blot result (Figure 10, (I) and (M) and Figure 13, Appendix).



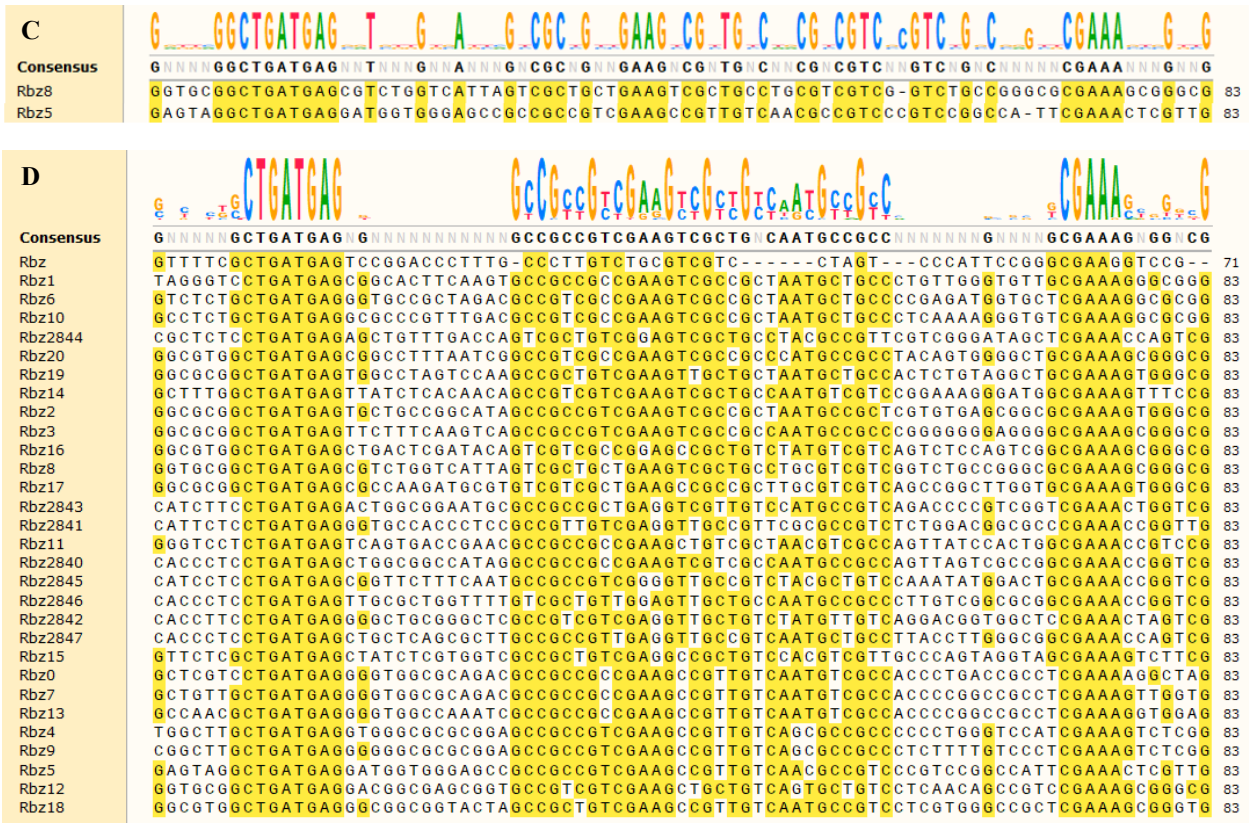


Figure 8. Ribozyme sequence alignments

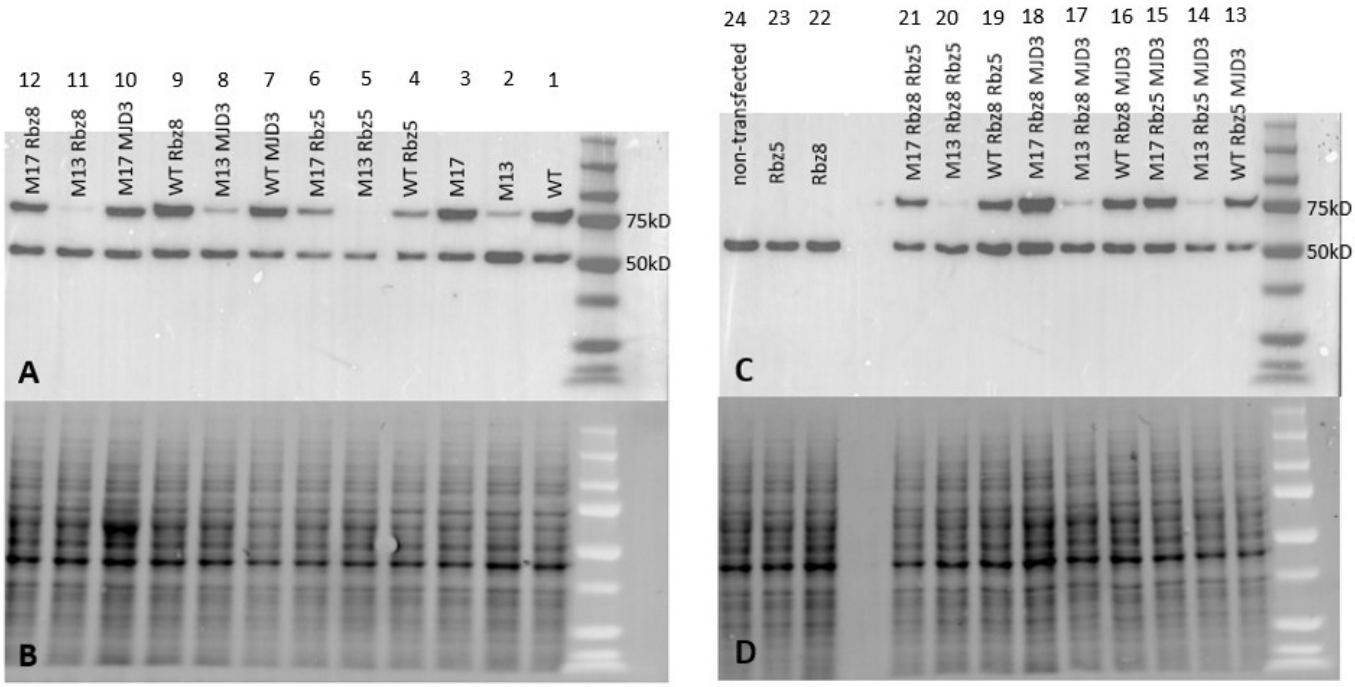
Every group of ribozymes designated in table 3 was aligned and the sequences in each group are lined in the order of similarity to the consensus sequence. Image A shows the efficient non-selective ribozymes, group B are the non-efficient non-selective ribozymes and group C only contains the two efficient and selective ribozymes. Image D shows the consensus sequences of all 29 tested ribozymes. All panels show the conservation of the catalytic core of the ribozymes and a great degree of commonality in the OBS regions. The binding arms of the ribozymes can bind to sequences of the PABPN1 mRNA father from the trinucleotide sequence, and are therefore different, but since mRNA is flexible the OBS can still bind the repeat sequences. However, similarities in sequences does do seem to be directly related to efficiency of selectivity. The order of the ribozymes in each panel confers their similarity to the consensus, which bears little information as to why the ribozymes function differently from one another. The panels were generated using multiple sequence alignment tool CLUSTALW [57].

3.3 Rbz5 and Rbz8 successfully knocked down PABPN1 mutant protein levels

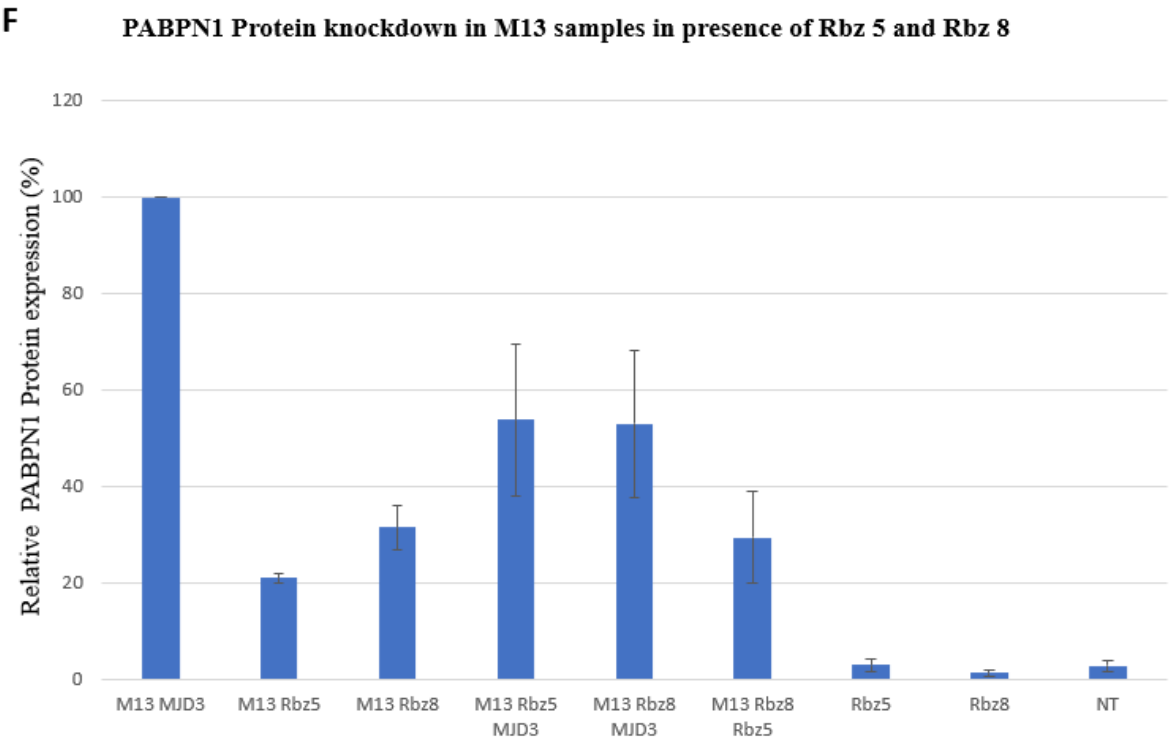
The ribozymes that selectively inhibit expression of PABPN1 mutant proteins can be valuable therapeutic tools since the disease-causing agents are believed to be the mutant proteins [58]. Figure 9 demonstrate how the mutant PABPN1 proteins were knocked down using Rbz5 and Rbz8.

PABPN1 is the polyadenylate-binding nuclear protein that is responsible for post transcriptional modification of messenger mRNA and is an integral protein in cells of most tissues [27, 58]. This means that the HEK293 cells used in this project's experiments have the endogenous PABPN1

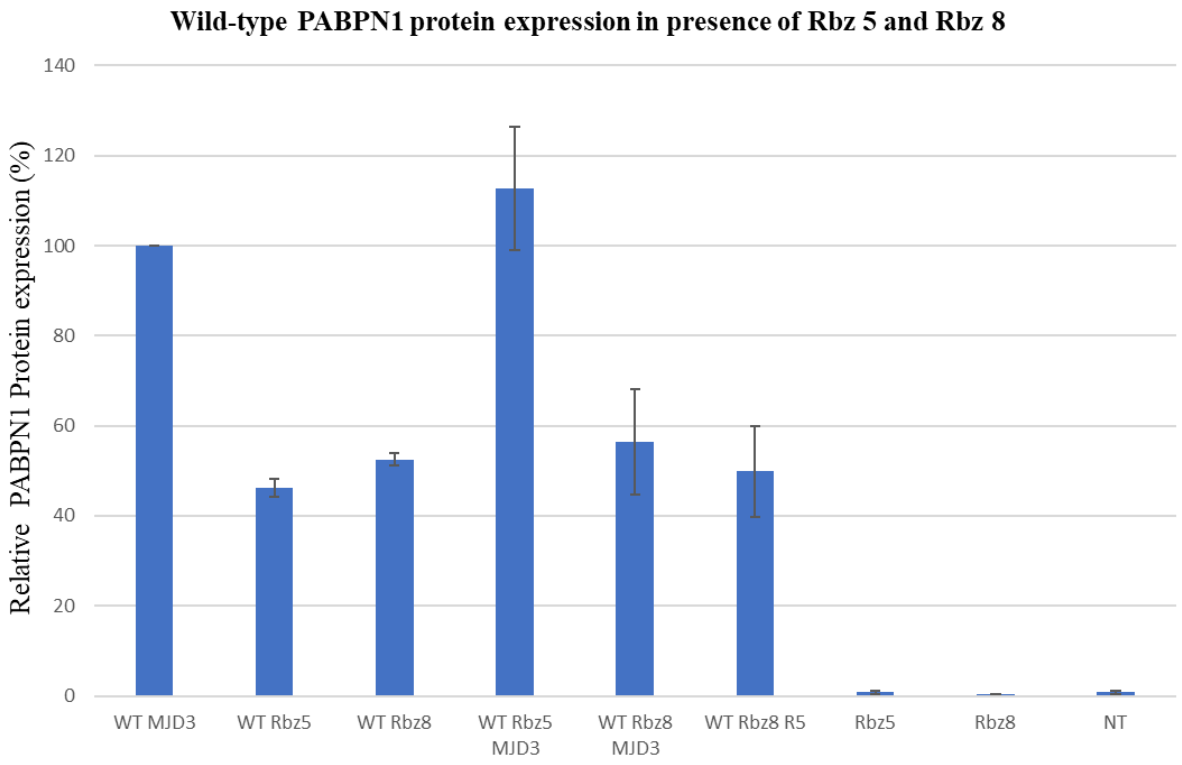
proteins (bands at 50 kD), which will be detected on all protein blots. Figure 9 ((A)s and (C)s) show a clear reduction in the overexpressed PABPN1 mutant proteins (bands at 75 kD). Both ribozymes (Rbz5 and Rbz8) show effective knock down of the mutant 13-Ala (Figure 9 (A), columns 5 and 6), and to a lesser extent of mutant 17-Ala (Figure 9 (A), columns 11 and 12) of PABPN1 protein. The combination of the ribozymes (Figure 9 (B), columns 20 and 21) is also effective in inhibiting the production of the mutant proteins.



F


PABPN1 Protein knockdown in M13 samples in presence of Rbz 5 and Rbz 8
Relative PABPN1 Protein expression (%)
120
100
80
60
40
20
0
M13 MJD3
M13 Rbz5
M13 Rbz8
M13 Rbz5 MJD3
M13 Rbz8 MJD3
M13 Rbz8 Rbz5
Rbz5
Rbz8
NT

E


Wild-type PABPN1 protein expression in presence of Rbz 5 and Rbz 8
Relative PABPN1 Protein expression (%)
140
120
100
80
60
40
20
0
WT MJD3
WT Rbz5
WT Rbz8
WT Rbz5 MJD3
WT Rbz8 MJD3
WT Rbz8 R5
Rbz5
Rbz8
NT

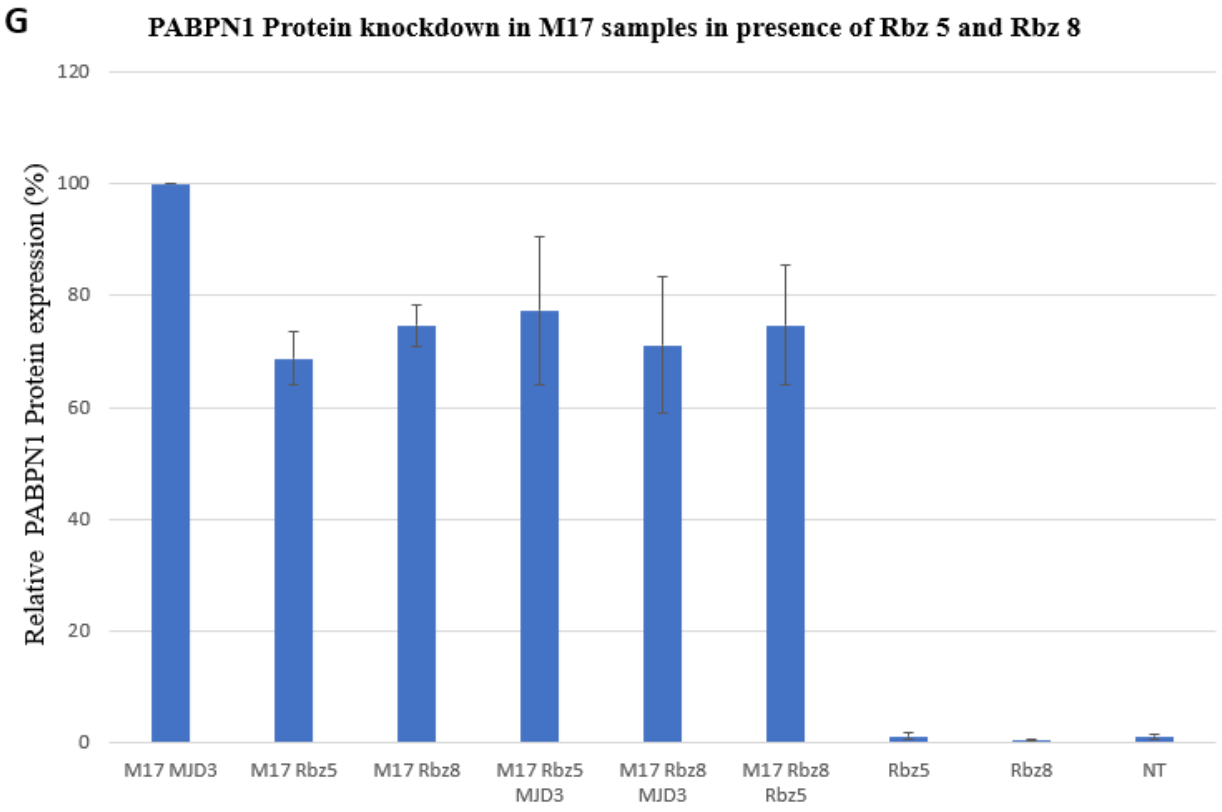


Figure 9. Effective knockdown of mutant 13-ala PABPN1 protein levels using Rbz5 and Rbz8

Western blots show protein knockdown using Rbz5 and Rbz8 in HEK293 cells when co-transfected with PABPN1 genes (WT, Mutant 13-Ala, and Mutant 17-Ala) and the two ribozymes, Rbz5 and Rbz8. Blots (A) and (C) depict overexpressed PABPN1-GFP fusion protein at 75 kD and the endogenous PABPN1 at 50 kD. The lanes are numbered from 1 to 24. Lanes 1 to 3 contain the proteins from cells that were transfected with a plasmid coding wild type PABPN1 fused to GFP or either of two mutants fused to GFP without any ribozymes. Lanes 4 to 6 shows protein expressions in presence of Rbz5 with a clear reduction in M13 and M17 protein bands. Lanes 7, 8, and 10 show the expression level in presence of the control ribozyme (MJD3) and 9, 11, and 12 show the effect of Rbz8 on PABPN1 protein expression with a decrease in M13 protein again. In blot (C) wells 13 to 21 each had PABPN1 proteins expressed in presence of two ribozymes at a time. 13 to 15 had Rbz5 and the control (MJD3) ribozymes and 16 to 18 contained Rbz8 and the control. Lanes 19, 20, and 21 had a combination of Rbz5 and Rbz8 with the aim of investigating the possibility of an additive effect. Blots (B) and (D) are total protein content of each well placed directly under the corresponding bands on (A) and (C). These samples were collected form HEK293 cells 24hrs post-transfection. The total protein content of each sample was used to normalize the quantities of each band in blot analysis. The experiment was repeated and generated similar results. The graphs (E), (F) and (G)

show the relative protein expression inhibition for PABPN1 wild-type (WT), mutant-13 Ala (M13) and mutant-17 Ala (M17) respectively. The graphs were generated using the quantification data from two sets of blots. MJD3 is the control ribozyme and co-transfection with this ribozyme is taken to be the maximum level of protein expression for normalization of each set of samples. NT represents the non-transfected samples and Rbz5, Rbz8 are the samples from cells that were transfected with the ribozymes alone (no overexpression of the PABPN1 gene). Error bars for each sample are indicated on the graphs.

The next step was quantifying these bands on the blots. ImageJ software [56] was used to quantify each band. The values were then subtracted from each blot's background and divided by the obtained measurements from the corresponding total protein content column. The averaged numbers from the two sets were then normalized to the designated controls (WT MJD3, M13 MJD3 and M17 MJD3). The MJD3 plasmid carries a ribozyme designed to target mutant *ataxin-3* gene with a polyglutamine (CAG) expansion. This means the ribozyme will not target the PABPN1 mRNA cleavage site, nor its GCG alanine repeats. MJD3 ribozyme was used as the control instead of an empty plasmid (containing no ribozyme). This was because transfection with an empty plasmid is not comparable to a transfection that leads to transcription of a ribozyme. The MJD3-carrying control plasmid accounts for the potential effects that expression of genes from two separate plasmids can have on the cells and the overall level of expression.

Quantification of blot results depicts the percent decrease in the relative (relative to the samples that have the Rz that does not target PABPN1) level of proteins (Figure 9 E, F, and G). The results show an overall decrease in protein levels in presence of the two ribozymes: Rbz5 and Rbz8 (Figure 9 E, F, and G) and a decrease (almost comparable to that of Rbz5) in expression of PABPN1 mutant-13Ala when cells were co-transfected with Rbz5 and Rbz8 (Figure 9F). Rbz5 and Rbz8 inhibited PABPN1 mutant-13Ala expression by ~80% and ~70% respectively (Figure 9 F). The effect of the ribozymes on mutant-17Ala seems less prominent with a reduction in protein levels of around 30% for both Rbz5 and Rbz8 (Figure 9 G). The reduction in the expression of the wild-type protein is around 50% for either of the ribozymes (Figure 9 E) it is thus targeted to a lesser extent than the mutant-13Ala.

The blot results obtained from Rbz6 and Rbz7 (efficient and non-selective) and Rbz2844 (non-efficient non-selective) samples show results representative of the groups the ribozymes are categorized in (Figure 12 and 13, Appendix). Co-transfection with Rbz6 results in an almost complete elimination of wild type and mutant 13-Ala proteins (around 95%) and ~70% decrease in mutant 17-Ala proteins. Rbz7 generated a similar outcome with a reduction of over ~70% in wild type protein and ~56% in mutant 17-Ala as well as a 36% decrease in mutant 13-Ala protein levels (Figure 12, Appendix). Rbz2844 which is one of the ribozymes in the non-efficient non-selective category, however, shows less than 10% cleavage in wild type and mutant 13-Ala PABPN1 and no reduction in mutant 17-Ala protein level (Figure 13, Appendix).

3.4 Quantifying mRNA transcript levels using RT-qPCR

To measure the amount of PABPN1 transcripts in the cells post-transfection, quantitative RT-PCR was performed on the cDNA which was reverse transcribed from the RNA extracts of these cells. The RNA was extracted from wells which were identical to the ones used for protein extraction. The goal was to compare the relative quantity of the mRNA transcripts in the presence or absence of the ribozymes. The percent reduction in the number of transcripts in the cells would be the direct result of the cleavage of the transcripts by the ribozymes and is therefore a measure of the ribozymes' performance. The RT-qPCR was performed on the ribozyme samples that showed efficient reduction in protein expression on blots (Table 4, Appendix). Therefore, nineteen ribozymes were tested by quantitative RT-PCR (Figures 16 and 14, Appendix). Many rounds of RT-qPCR were performed, and the results were carefully recorded (Figures 16 and 14, Appendix and Table 4, Appendix) however, the results could not be interpreted in a meaningful way in most cases (Figure 16, Appendix) as clear patterns could not be detected and unexpected discrepancies were often observed.

Chapter 4: Discussion

4.1 Hammerhead ribozymes can be designed to target PABPN1 transcripts

PABPN1 protein activity is a crucial part of the physiology of many cells, including muscle cells and neurons [27], whereas aggregation of the mutant proteins is known to be responsible for the development of the disease known as oculopharyngeal muscular dystrophy (OPMD) [59]. Presently, there is no cure for OPMD, and treatments mostly aim to alleviate the symptoms rather than targeting the cause. This means that the patients can never be fully relieved of the symptoms of the disease, and although the condition does not usually reduce life expectancy greatly, it progressively lowers the quality of life of the patients [20].

Different methods of gene therapy are being studied with the purpose of targeting the disease at the DNA, RNA and protein levels [43, 60, 61]. The condition is monogenic, and the mutation causing it is an expansion of alanine repeats, which in theory make the transcript a good target for ribozymes such as those designed using *TriCleaver*. Using ribozymes gives a clear advantage over certain gene editing methods, such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), which use guide RNAs. This is because one cannot design a guide RNA that distinguishes between different numbers of repeats. The guide RNA is relatively short (up to 20 nucleotides) and cannot contain many repeats. A study by Ran et al. has shown that even if the guide RNA of a longer sequence is used, it will be trimmed to 20 nucleotides when being processed *in vivo* [62] .

This project aimed to identify the ribozymes (from a pool of designed ribozymes) that can target the mutant PABPN1 transcripts selectively and hence, have minimal effects on expression of the wild-type protein.

Among the 29 ribozymes designed by the algorithm *Tricleaver*, and tested through this project, Rbz5 and Rbz8 were found to be efficient and relatively selective for mutant transcripts. The preliminary results from this project are to be used to optimize the designs of the selected ribozymes to increase the selectivity and effectiveness. Rbz5 was shown to reduce the mutant-13 Ala transcript levels to 20% while affecting the wild type only 50%. This shows that the design was successful and effective in cleaving this mutant mRNA and relatively selective in targeting the mutant transcript. Rbz8 has produced good results by reducing the mutant-13 Ala to around 30% while the wild-type PABPN1 protein is reduced to 50%.

4.2 Limitations, discrepancies, and proposed modifications for improved results

As discussed in the results section, RT-qPCR of all samples were done in triplicates (Figure 16 and table 4, Appendix), however these results could not be meaningfully analyzed. At times the results show a higher level of transcription than expected. An example of this would be when the level of transcription in presence of the ribozyme(s) is much higher than in the absence of the

ribozyme. This implies that the ribozyme enhances the transcription of the gene instead of cleaving and eliminating the transcript. These odd results could be associated with problems with the RT-qPCR experiment. Although qPCR results had acceptable error bars for our triplicates, several rounds of qPCR showed inconsistent results. Attempts to optimize the assay, by recalibrating the machine, changing plate readers, calibrating pipettes, and repeating the experiments failed to resolve the issue. This led us to speculate that part of the issue may lie within the cells and the inherent properties of both the HEK293 cells and PABPN1 gene and transcript. Different cellular pathways, expression of housekeeping genes, and how close the cell is to apoptosis when the RNA and protein are extracted [63, 64] are all factors that can affect the level of transcripts in the cells.

There are several hypotheses which can be tested to explain the inconsistent RT-qPCR results.

1. The half-life of PABPN1 transcript in HEK293 cells is not well-established but it is known to be unstable in C2C12 and NIH/3T3 cells [65]. From this, it is reasonable to infer that the transcript may be unstable in HEK293 cells as well. Proteins and RNA in this project were either extracted after 48hrs or 24hrs post-transfection, which means the level of mRNA transcripts may not have been comparable to the level of proteins produced by this mRNA in the cells anymore. Amplifying the extracted RNA then would not be representative of the protein levels present in cells. The instability of wild -type PABPN1 mRNA in physiological conditions [65] raises the question of the stability of the mutant transcripts as well, which can mean an inconsistency in the levels of mRNA of the wild-type compared to mutant-13 Ala and mutant-17 Ala transcripts. To investigate the relevance of the transcript half-lives and stability, RNA can be extracted at different time intervals following treatment with actinomycin D (which stops transcription) to estimate the optimized time of extraction following arrest of mRNA production for acceptable qPCR results.

2. Quantity and quality of extracted RNA can be affected by storage time. Although all RT-qPCR experiments were done within a short time after extraction, the time interval was not fixed meaning even if the RNA decay rate in all samples is taken to be the same, the storage time of samples from cells with different ribozymes were not controlled for. The results may be improved if the extracted mRNA is reverse-transcribed into cDNA shortly after extraction.

3. Another factor that might affect the level of expression of PABPN1 genes is the fact that the overexpression of mutant PABPN1 creates aggregates in the nucleus of the cells that can result in cell death [66]. Hence the cells that were transfected with the mutant plasmids expressed mutant protein that might have formed aggregates inside the nucleus of the cells and could change the global transcription patterns by interfering with the expression of the essential genes or resulting in cell apoptosis. A shorter exposure time may minimize the effect of the toxic proteins, hence RNA extraction a short time after transfection may help. Although in that case, it is also possible that the limited time is not enough to fully assess the effect of the ribozymes on their targets.

4. In comparative real-time qPCR, we assume that the relative levels of expression of the target gene (PABPN1) compared to the endogenous control gene (RNA polymerase II) does not vary, except when we treat with ribozymes targeting PABPN1 [67]. This assumption can be another reason for the discrepancies in the fold differences in the results as the level of expression of PABPN1 may vary in conditions different than that of the internal control. This is especially important when we look at the results from PABPN1 mutant genes as the abnormal transcripts and the resulting proteins may affect the level of RNA Pol II transcript and subsequently the translated

protein, given that PABPN1 is a polyA binding protein known to stabilize mRNAs. These mutants could be changing the stability of their own mRNA as well which can make the results more confusing. PABPN1 controls the alternative polyadenylation and tri-nucleotide expansion of proteins. PABPN1 mutant 17-Ala has been shown to cause shortening of the 3'-UTR (a region responsible for post-translational modifications) globally [68]. This means that the reduction on the level of PABPN1 caused by the ribozymes can influence the levels of the wild type and mutant PABPN1 as well as other housekeeping genes within the cells. Hence, the use of multiple reference genes, meaning two or three endogenous controls, may increase the accuracy of the RT-qPCR results.

5. When cells are transfected with certain plasmids, they may take up different amounts of each plasmid. In the experiments performed for this project, we only had GFP protein markers on the gene-expressing plasmids and no detectable markers on the ribozyme plasmids. This means that the efficiency of transfection could be determined for the overexpressed PAPBN1 wild-type, mutant 13-Ala and mutant 17-Ala proteins, but not for the ribozyme carrying plasmids. As a result, the quantity of ribozymes, which are available to catalyze mRNA cleavages, in cells is unclear. This issue can be addressed by including primers that amplify the ribozymes during RT-qPCR as well as the PABPN1-carrying genes. However, this would not take into account the fact that plasmids may not always be co-transfected in the same cells. The theoretical worst-case example of 50% efficiency transfection would be 50% of cells taking up the PABPN1 transgene and the other 50% having the ribozyme plasmid, so 0% cleavage would be observed. More realistically, in such a case we could expect that 25% of cells would be co-transfected, thus showing only half the effect that the ribozyme truly has on cells. Hence, variations from assay to assay of this co-transfection efficiency could significantly impact the results.

6. Another less probable but not impossible reason for the discrepancies between the blot and the qPCR results could be from amplification of transcripts that may not have been completely degraded at the time of RNA extraction. When the ribozyme cleaves the mRNA, each piece will either have the 3' polyA tail or the 5'cap. The assumption of this project was that these mRNA molecules would be rapidly degraded [69]. However, if these molecules are still present when reverse transcription prior to RT-qPCR is performed, the results will be affected and cannot explain the blot results. This might be a possibility as these mRNA molecules that have been cleaved cannot be translated into proteins but can be reverse transcribed into cDNA which is stable and is quantified through qPCR. The primers that are used for qPCR produce an amplicon which is quite small (107 base pairs). This means that the primers can possibly bind and amplify the cDNA that does not cover the whole span of the gene but still contains the sequence the primers can bind to, and the probe will detect. Using primer pairs that overlap each cleavage site during qPCR would ensure that the cleavage is detected. Another modification can be during cDNA synthesis, the product of which is used in qPCR. This project used the SuperScript IV VILO kit (Invitrogen) for cDNA synthesis that contains short random primers which can bind anywhere on any mRNA present in RNA extracts. Using specific primers that will only bind *PABPN1* mRNA during cDNA synthesis, could help in optimizing the qPCR results.

4.3 Conclusion and future directions

This work served as an evaluation platform for the *TriCleaver* algorithm [48] by assessing the generated sequences. It is important to note that many of the designs were found to be effective in cleaving the mRNA but did not meet the selectivity criterion.

Two ribozymes, Rbz5 and Rbz8, showed did exhibit a marked level of efficiency and selectivity in knocking down mutant PABPN1 gene expression, at the protein level. However, given some problems associated with inter-sample variations, especially at the mRNA level, the results should be interpreted cautiously. Nevertheless, the results of these two ribozymes do provide indications that the algorithm has the potential to produce designs that can selectively target the mRNA of various trinucleotide expansion diseases.

The end goal of this algorithm is to design ribozymes that can selectively and effectively cleave the mutant mRNAs associated with different nucleotide expansion diseases; it will be improved based on the results obtained in this project, to generate new sequences with a greater chance of selectively targeting the diseased transcripts.

The obtained results provide insight into how new ribozymes can be designed. The sequence of the selective ribozymes as well as the highly effective ribozymes, and even the inactive ribozymes, can be used as a basis for creation of other potentially selective ribozymes. Indeed, ribozymes that cleave both wild type and mutant versions compared to those that cleave none of the targets could help inform us on the free energy differences between the inactive and active ribozymes that we should aim for.

Finally, the ribozymes (Rbz5 and Rbz8) that were found to selectively cleave the PABPN1 mutant genes can be tested in *C. elegans* and mouse models for further *in vivo* confirmation.

References

1. Haubold, B. and T. Wiehe, *How repetitive are genomes?* BMC Bioinformatics, 2006. **7**(1): p. 541.
2. de Koning, A.P., et al., *Repetitive elements may comprise over two-thirds of the human genome.* PLoS Genet, 2011. **7**(12): p. e1002384.
3. Paulson, H., *Repeat expansion diseases.* Handb Clin Neurol, 2018. **147**: p. 105-123.
4. Richard, G.-F., *The Startling Role of Mismatch Repair in Trinucleotide Repeat Expansions.* Cells, 2021. **10**(5): p. 1019.
5. Limprasert, P., et al., *Analysis of CAG repeat of the Machado-Joseph gene in human, chimpanzee and monkey populations: a variant nucleotide is associated with the number of CAG repeats.* Hum Mol Genet, 1996. **5**(2): p. 207-13.
6. McMurray, C.T., *Mechanisms of trinucleotide repeat instability during human development.* Nat Rev Genet, 2010. **11**(11): p. 786-99.
7. Wheeler, V.C. and V. Dion, *Modifiers of CAG/CTG Repeat Instability: Insights from Mammalian Models.* J Huntingtons Dis, 2021. **10**(1): p. 123-148.
8. Petruska, J., M.J. Hartenstine, and M.F. Goodman, *Analysis of Strand Slippage in DNA Polymerase Expansions of CAG/CTG Triplet Repeats Associated with Neurodegenerative Disease *.* Journal of Biological Chemistry, 1998. **273**(9): p. 5204-5210.
9. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
10. Häsler, J. and K. Strub, *Alu elements as regulators of gene expression.* Nucleic Acids Research, 2006. **34**(19): p. 5491-5497.
11. Kim, S., et al., *Structural Variation of Alu Element and Human Disease.* Genomics Inform, 2016. **14**(3): p. 70-77.
12. Price, A.L., E. Eskin, and P.A. Pevzner, *Whole-genome analysis of Alu repeat elements reveals complex evolutionary history.* Genome Res, 2004. **14**(11): p. 2245-52.
13. Larsen, P.A., et al., *Warning SINEs: Alu elements, evolution of the human brain, and the spectrum of neurological disease.* Chromosome Res, 2018. **26**(1-2): p. 93-111.
14. Kurosaki, T., et al., *The unstable CCTG repeat responsible for myotonic dystrophy type 2 originates from an AluSx element insertion into an early primate genome.* PLoS One, 2012. **7**(6): p. e38379.
15. Cummings, C.J. and H.Y. Zoghbi, *Trinucleotide Repeats: Mechanisms and Pathophysiology.* Annual Review of Genomics and Human Genetics, 2000. **1**(1): p. 281-328.
16. Jurka, J., et al., *Duplication, coclustering, and selection of human Alu retrotransposons.* Proceedings of the National Academy of Sciences, 2004. **101**(5): p. 1268-1272.
17. Shapiro, J.A. and R. von Sternberg, *Why repetitive DNA is essential to genome function.* Biol Rev Camb Philos Soc, 2005. **80**(2): p. 227-50.
18. Lu, J.Y., et al., *Genomic Repeats Categorize Genes with Distinct Functions for Orchestrated Regulation.* Cell Rep, 2020. **30**(10): p. 3296-3311.e5.
19. Malerba, A., et al., *PABPN1 gene therapy for oculopharyngeal muscular dystrophy.* Nat Commun, 2017. **8**: p. 14848.
20. Yamashita, S., *Recent Progress in Oculopharyngeal Muscular Dystrophy.* J Clin Med, 2021. **10**(7).
21. Abu-Baker, A. and G.A. Rouleau, *Oculopharyngeal muscular dystrophy: recent advances in the understanding of the molecular pathogenic mechanisms and treatment strategies.* Biochim Biophys Acta, 2007. **1772**(2): p. 173-85.
22. Youssof, S., *The relationship between physical symptoms and health-related quality of life in oculopharyngeal muscular dystrophy.* Muscle Nerve, 2016. **53**(5): p. 694-9.

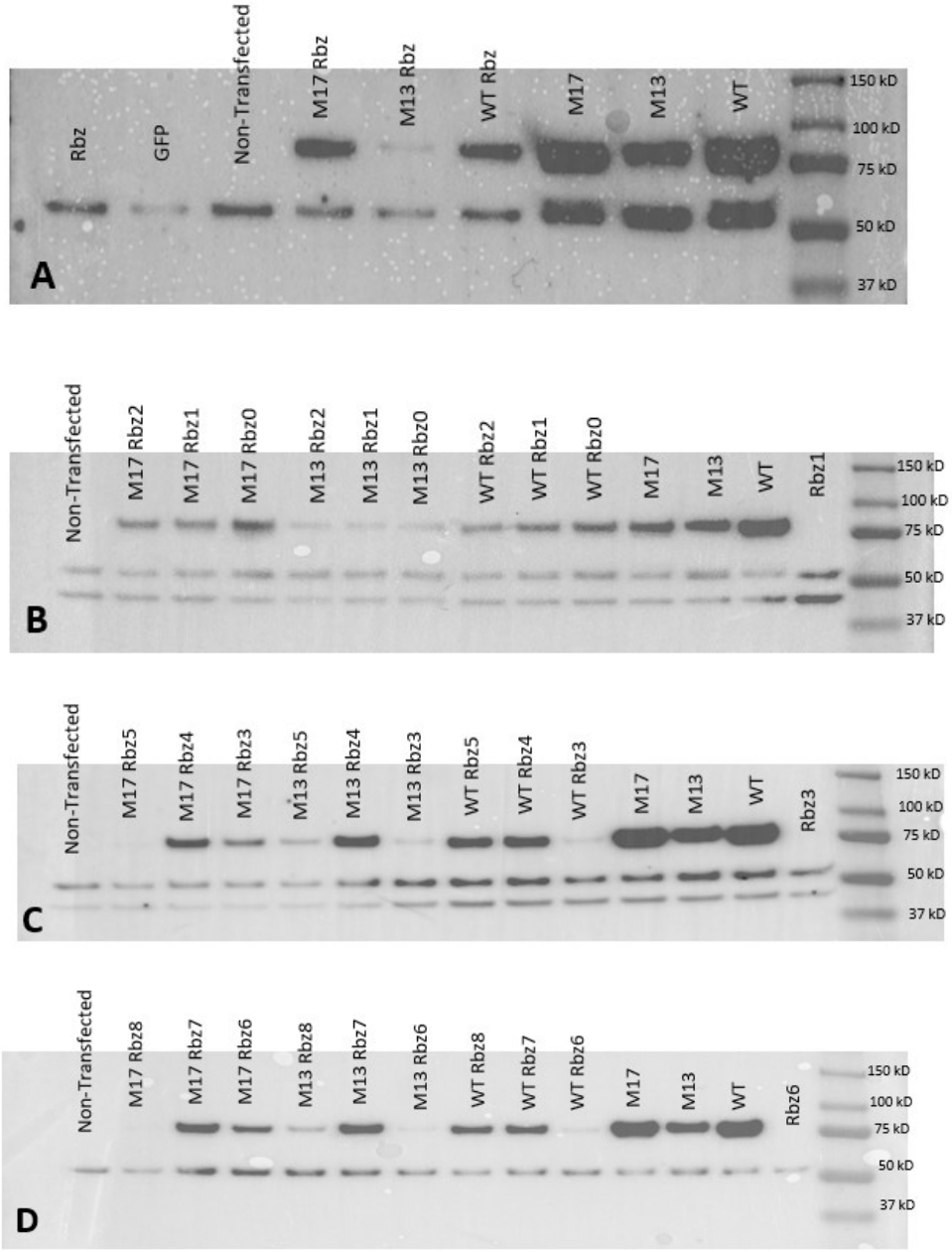
23. Tomé, F.M., et al., *Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy.* Neuromuscul Disord, 1997. **7 Suppl 1**: p. S63-9.
24. Corbeil-Girard, L.P., et al., *PABPN1 overexpression leads to upregulation of genes encoding nuclear proteins that are sequestered in oculopharyngeal muscular dystrophy nuclear inclusions.* Neurobiol Dis, 2005. **18**(3): p. 551-67.
25. Davies, J.E., S. Sarkar, and D.C. Rubinsztein, *Wild-type PABPN1 is anti-apoptotic and reduces toxicity of the oculopharyngeal muscular dystrophy mutation.* Hum Mol Genet, 2008. **17**(8): p. 1097-108.
26. Roth, F., et al., *Assessment of PABPN1 nuclear inclusions on a large cohort of patients and in a human xenograft model of oculopharyngeal muscular dystrophy.* Acta Neuropathol, 2022. **144**(6): p. 1157-1170.
27. Banerjee, A., et al., *PABPN1: molecular function and muscle disease.* Febs j, 2013. **280**(17): p. 4230-50.
28. Boutet, S.C., et al., *Alternative polyadenylation mediates microRNA regulation of muscle stem cell function.* Cell Stem Cell, 2012. **10**(3): p. 327-36.
29. Muniz, L., L. Davidson, and S. West, *Poly(A) Polymerase and the Nuclear Poly(A) Binding Protein, PABPN1, Coordinate the Splicing and Degradation of a Subset of Human Pre-mRNAs.* Mol Cell Biol, 2015. **35**(13): p. 2218-30.
30. Mayr, C. and D.P. Bartel, *Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.* Cell, 2009. **138**(4): p. 673-84.
31. Apponi, L.H., et al., *Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis.* Hum Mol Genet, 2010. **19**(6): p. 1058-65.
32. Becher, M.W., et al., *Oculopharyngeal Muscular Dystrophy in Hispanic New Mexicans.* JAMA, 2001. **286**(19): p. 2437-2440.
33. Blumen, S.C., et al., *Cognitive impairment and reduced life span of oculopharyngeal muscular dystrophy homozygotes.* Neurology, 2009. **73**(8): p. 596-601.
34. Forood, B., et al., *Formation of an extremely stable polyalanine beta-sheet macromolecule.* Biochem Biophys Res Commun, 1995. **211**(1): p. 7-13.
35. Bao, Y.P., et al., *Mammalian, yeast, bacterial, and chemical chaperones reduce aggregate formation and death in a cell model of oculopharyngeal muscular dystrophy.* J Biol Chem, 2002. **277**(14): p. 12263-9.
36. Davies, J.E., et al., *Cystamine suppresses polyalanine toxicity in a mouse model of oculopharyngeal muscular dystrophy.* Sci Transl Med, 2010. **2**(34): p. 34ra40.
37. Malerba, A., et al., *Pharmacological modulation of the ER stress response ameliorates oculopharyngeal muscular dystrophy.* Hum Mol Genet, 2019. **28**(10): p. 1694-1708.
38. Périé, S., et al., *Autologous myoblast transplantation for oculopharyngeal muscular dystrophy: a phase I/IIa clinical study.* Mol Ther, 2014. **22**(1): p. 219-25.
39. Tanner, N.K., *Ribozymes: the characteristics and properties of catalytic RNAs.* FEMS Microbiology Reviews, 1999. **23**(3): p. 257-275.
40. Walter, N.G. and D.R. Engelke, *Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs.* Biologist (London), 2002. **49**(5): p. 199-203.
41. Torres, R.A. and T.C. Bruice, *The Mechanism of Phosphodiester Hydrolysis: Near In-line Attack Conformations in the Hammerhead Ribozyme.* Journal of the American Chemical Society, 2000. **122**(5): p. 781-791.
42. Perreault, J., et al., *Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations.* PLOS Computational Biology, 2011. **7**(5): p. e1002031.
43. Abu-Baker, A., et al., *RNA-Based Therapy Utilizing Oculopharyngeal Muscular Dystrophy Transcript Knockdown and Replacement.* Molecular Therapy - Nucleic Acids, 2019. **15**: p. 12-25.

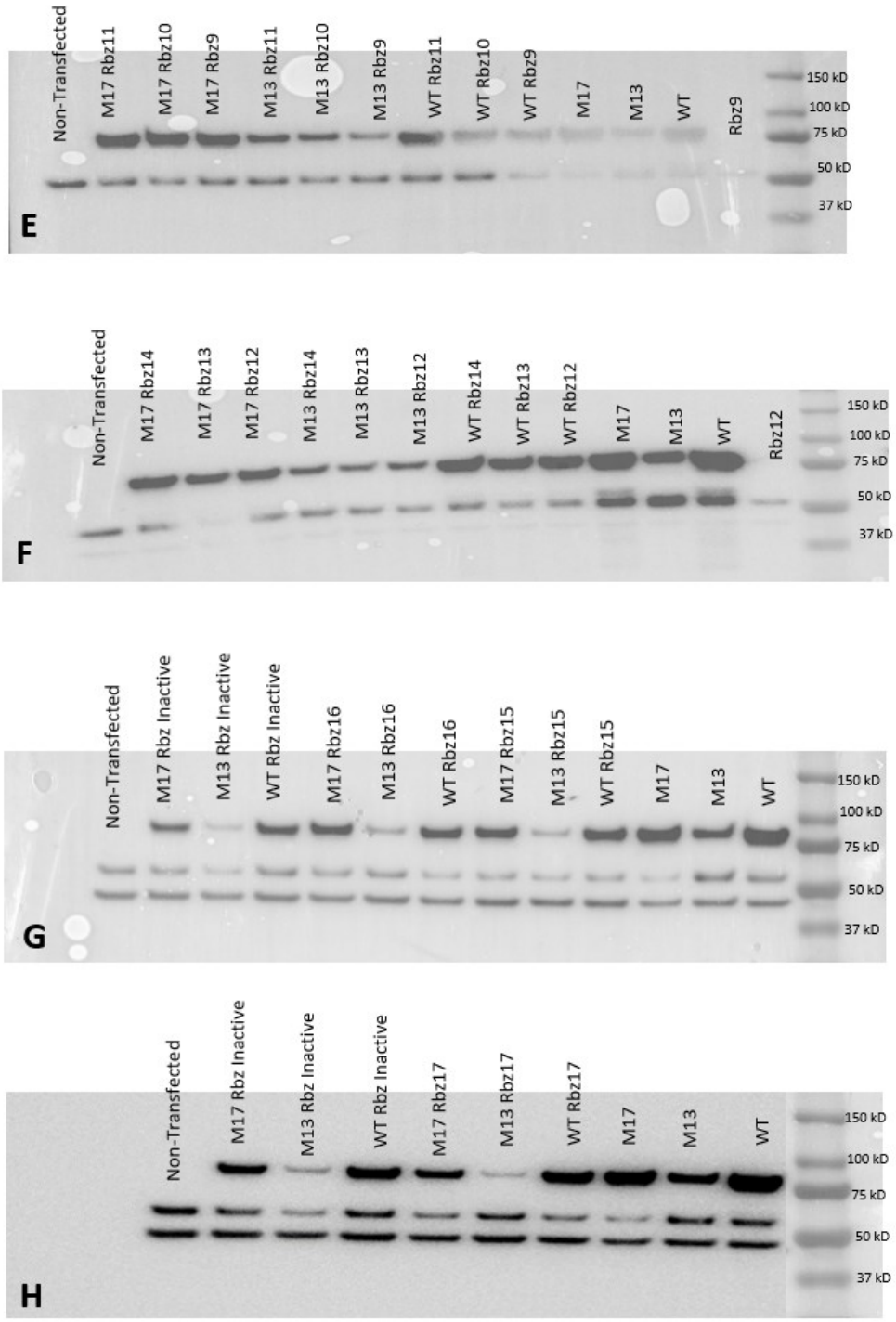
44. Hammann, C., et al., *The ubiquitous hammerhead ribozyme.* Rna, 2012. **18**(5): p. 871-85.
45. Luzi, E., F. Eckstein, and G. Barsacchi, *The newt ribozyme is part of a riboprotein complex.* Proceedings of the National Academy of Sciences, 1997. **94**(18): p. 9711-9716.
46. Huang, X., et al., *Intracellular selection of trans-cleaving hammerhead ribozymes.* Nucleic Acids Res, 2019. **47**(5): p. 2514-2522.
47. Penchovsky, R. and G.T. Kostova, *Computational selection and experimental validation of allosteric ribozymes that sense a specific sequence of human telomerase reverse transcriptase mRNAs as universal anticancer therapy agents.* Nucleic Acid Ther, 2013. **23**(6): p. 408-17.
48. Kamel, N., et al. *Evolutionary Design and Experimental Evaluation of Selective Hammerhead Ribozymes*. in *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*. 2020.
49. Lorenz, R., et al. *ViennaRNA Package 2.0*. Algorithms for molecular biology : AMB, 2011. **6**, 26 DOI: 10.1186/1748-7188-6-26.
50. Varani, G. and W.H. McClain, *The G·U wobble base pair.* EMBO reports, 2000. **1**(1): p. 18-23.
51. Masquida, B. and E. Westhof, *On the wobble GoU and related pairs.* Rna, 2000. **6**(1): p. 9-15.
52. Kerpedjiev, P., S. Hammer, and I.L. Hofacker, *Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams.* Bioinformatics, 2015. **31**(20): p. 3377-3379.
53. Nawrot, B., et al., *Efficient inhibition of beta-secretase gene expression in HEK293 cells by tRNAVal-driven and CTE-helicase associated hammerhead ribozymes.* Eur J Biochem, 2003. **270**(19): p. 3962-70.
54. Abu-Baker, A., et al., *Lithium chloride attenuates cell death in oculopharyngeal muscular dystrophy by perturbing Wnt/β-catenin pathway.* Cell Death & Disease, 2013. **4**(10): p. e821-e821.
55. Bustin, S.A., et al., *The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments.* Clinical Chemistry, 2009. **55**(4): p. 611-622.
56. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nature Methods, 2012. **9**(7): p. 671-675.
57. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Research, 1994. **22**(22): p. 4673-4680.
58. Klein, P., et al., *Nuclear poly(A)-binding protein aggregates misplace a pre-mRNA outside of SC35 speckle causing its abnormal splicing.* Nucleic Acids Res, 2016. **44**(22): p. 10929-10945.
59. Tomé, F.M. and M. Fardeau, *Nuclear inclusions in oculopharyngeal dystrophy.* Acta Neuropathol, 1980. **49**(1): p. 85-7.
60. Abu-Baker, A., et al., *Cytoplasmic targeting of mutant poly(A)-binding protein nuclear 1 suppresses protein aggregation and toxicity in oculopharyngeal muscular dystrophy.* Traffic, 2005. **6**(9): p. 766-79.
61. Robinson, D.O., et al., *Oculopharyngeal muscular dystrophy: a point mutation which mimics the effect of the PABPN1 gene triplet repeat expansion mutation.* J Med Genet, 2006. **43**(5): p. e23.
62. Ran, F.A., et al., *Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.* Cell, 2013. **154**(6): p. 1380-9.
63. Uhlén, M., et al., *Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
64. Singh, K.P., et al., *Mechanisms and Measurement of Changes in Gene Expression.* Biol Res Nurs, 2018. **20**(4): p. 369-382.
65. Phillips, B.L., et al., *Post-transcriptional regulation of Pabpn1 by the RNA binding protein HuR.* Nucleic Acids Res, 2018. **46**(15): p. 7643-7661.

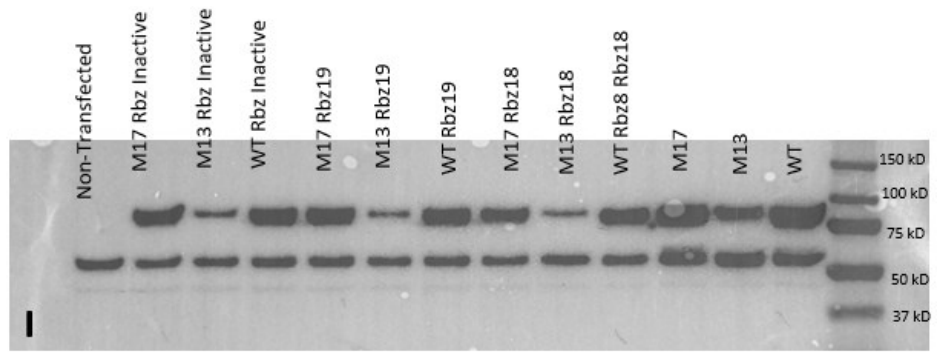
66. Messaed, C., et al., *Soluble expanded PABPN1 promotes cell death in oculopharyngeal muscular dystrophy.* Neurobiol Dis, 2007. **26**(3): p. 546-57.
67. Schmittgen, T.D. and K.J. Livak, *Analyzing real-time PCR data by the comparative C(T) method.* Nat Protoc, 2008. **3**(6): p. 1101-8.
68. Jenal, M., et al., *The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites.* Cell, 2012. **149**(3): p. 538-53.
69. Tharun, S. and R. Parker, *Targeting an mRNA for Decapping: Displacement of Translation Factors and Association of the Lsm1p–7p Complex on Deadenylated Yeast mRNAs.* Molecular Cell, 2001. **8**(5): p. 1075-1083.

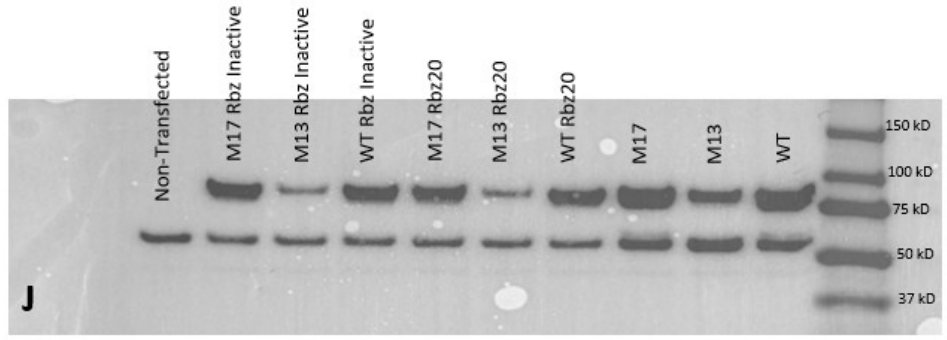
Appendix

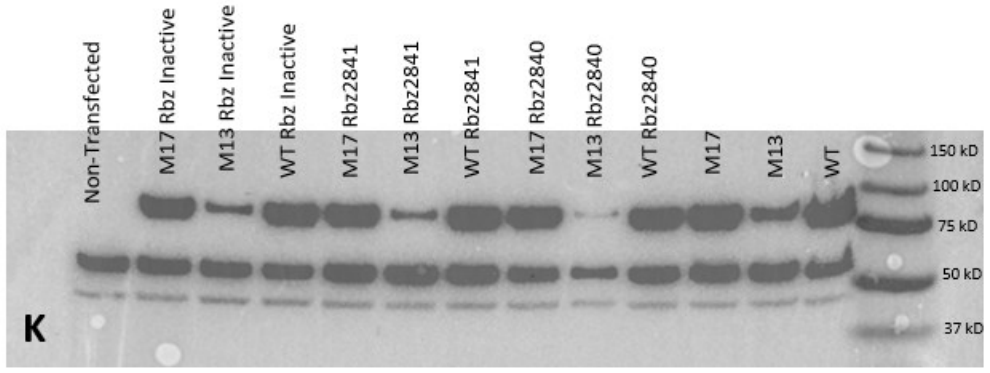
A1. Blot images of all ribozymes tested

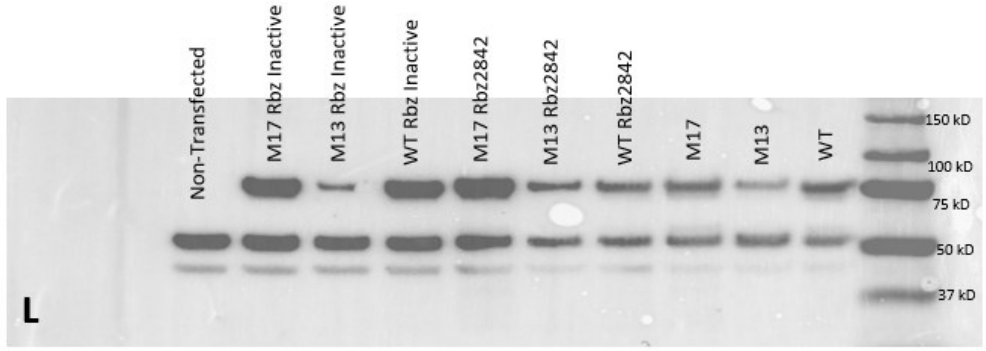



E
Non-Transfected
M17 Rbz11
M17 Rbz10
M17 Rbz9
M13 Rbz11
M13 Rbz10
M13 Rbz9
WT Rbz11
WT Rbz10
WT Rbz9
M17
M13
WT
Rbz9
150 kD
100 kD
75 kD
50 kD
37 kD
F
Non-Transfected
M17 Rbz14
M17 Rbz13
M17 Rbz12
M13 Rbz14
M13 Rbz13
M13 Rbz12
WT Rbz14
WT Rbz13
WT Rbz12
M17
M13
WT
Rbz12
150 kD
100 kD
75 kD
50 kD
37 kD
G
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz16
M13 Rbz16
WT Rbz16
M17 Rbz15
M13 Rbz15
WT Rbz15
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD
H
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz17
M13 Rbz17
WT Rbz17
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD


I
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz19
M13 Rbz19
WT Rbz19
M17 Rbz18
M13 Rbz18
WT Rbz8 Rbz18
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD


J
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz20
M13 Rbz20
WT Rbz20
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD


K
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz2841
M13 Rbz2841
WT Rbz2841
M17 Rbz2840
M13 Rbz2840
WT Rbz2840
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD


L
Non-Transfected
M17 Rbz Inactive
M13 Rbz Inactive
WT Rbz Inactive
M17 Rbz2842
M13 Rbz2842
WT Rbz2842
M17
M13
WT
150 kD
100 kD
75 kD
50 kD
37 kD

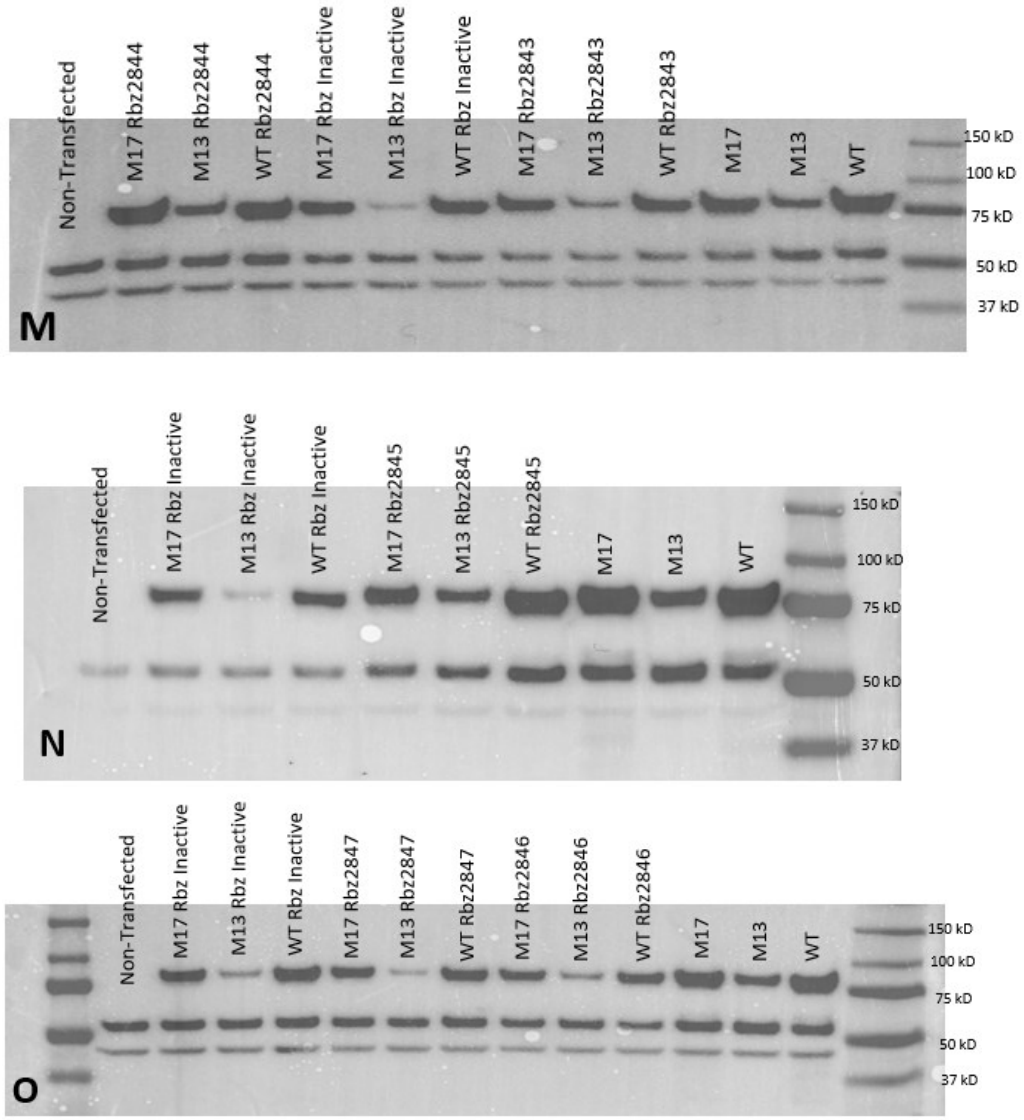


Figure 10. Blot results of all twenty-nine ribozymes

Blots are presented here in image A to O. Every column is labeled with the plasmids that were used to transfect the cells from which the proteins in that column were extracted. WT represents the plasmid that carried the wild type PABPN1 gene with 10 alanine repeats. M13 represents the plasmid that carried the mutant PABPN1 gene with 13 alanine repeats, and M17 represents the plasmid that carried the mutant PABPN1 gene with 17 alanine repeats. Rbz represents the plasmids carrying the ribozymes are numbered (Rbz, Rbz0, …, Rbz20 & Rbz2840, …, Rbz2847). Non-transfected means the cells were not transfected with any genes (negative control) and the inactive ribozyme is used as control since it cannot target PABPN1 transcripts. The bands on the blots from the top are the overexpressed PABPN1 proteins, the endogenous PABPN1, and Actin (initially used as the control to determine if loadings were equal in wells). Actin bands were not used to normalize the protein loads, instead the total protein content, samples of which are shown in the results section, were used for normalizations. GFP (only in image A) is green fluorescent protein used as a positive control. As can be seen on the blots,

M13 protein bands were always less pronounced as M13 protein was always expressed in smaller amounts compared to WT and M17 merely because the promoter on the plasmid that expressed M13 was different from (and less efficient than) the promoters on WT and M17 plasmids. This does not affect the results as M13 samples are only compared to other M13 samples and never to WT or M17.

A2. Blot Calculations for Rbz8 and Rbz5 samples

A		Blot 1 set 1 Rbz5 and Rbz8 combination set				(Empty box - band intensity mean)	Total protein content				(Empty box for total protein - mean total protein)	(Empty -mean)/(Empty total protein- mean total protein)
Sample content	#	area	Mean	Min	Max	(221.912 - mean)	area	Mean	Min	Max	(175.933 - mean)	(221.912 - mean)/(175.933 - mean)
WT PABPN1	1	0.019	150.792	93	230	71.12	11640	130.278	0	181	45.655	1.557770233
M13 PABPN1	2	0.019	203.082	159	226	18.83	11640	121.227	0	179	54.706	0.344203561
M17 PABPN1	3	0.019	154.122	94	218	67.79	11640	130.242	0	178	45.691	1.48366199
WT + Rbz5	4	0.019	193.755	122	247	28.157	11640	132.141	0	182	43.792	0.642971319
M13 + Rbz5	5	0.019	218.74	212	248	3.172	11640	125.833	9	177	50.1	0.063313373
M17 + Rbz5	6	0.019	192.087	112	225	29.825	11640	131.888	15	179	44.045	0.677148371
WT + MJD3 (control)	7	0.019	162.575	93	219	59.337	11640	138.697	15	181	37.236	1.593538511
M13 + MJD3 (control)	8	0.019	202.697	160	235	19.215	11640	122.628	0	179	53.305	0.360472751
M17 + MJD3 (control)	9	0.019	150.266	96	224	71.646	11640	115.473	0	181	60.46	1.185014886
WT + Rbz8	10	0.019	164.785	94	223	57.127	11640	101.452	0	178	74.481	0.76700098
M13 + Rbz8	11	0.019	218.521	201	232	3.391	11640	115.929	0	181	60.004	0.056512899
M17 + Rbz8	12	0.019	171.765	98	225	50.147	11640	111.452	0	179	64.481	0.777701959
Empty box	13	0.019	221.912	216	230		11640	175.933	162	186		

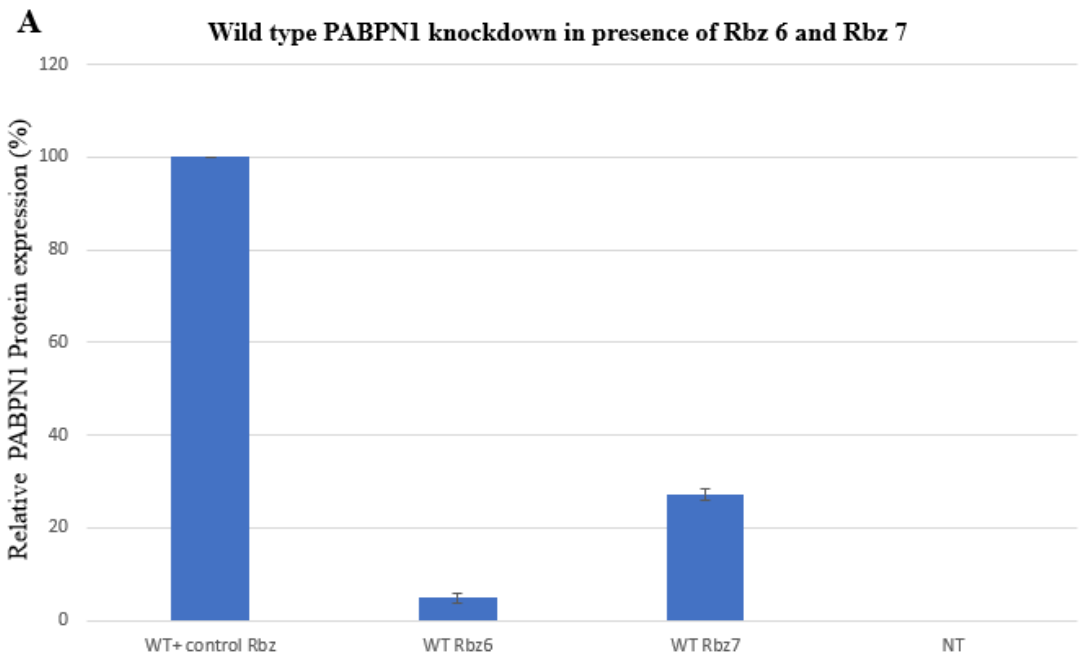
B		Blot 2 set 1 Rbz5 and Rbz8 combination set				(Empty box - band intensity mean)	Total protein content				(Empty box for total protein - mean total protein)	(Empty -mean)/(Empty total protein- mean total protein)
	#	area	Mean	Min	Max	(220.556 - mean)	area	Mean	Min	Max	(157.135 - mean)	(220.556 - mean)/(157.135 - mean)
WT + Rbz5 + MJD3 (control)	1	1950	172.213	93	220	48.343	0.145	135.903	19	172	21.232	2.276893369
M13 + Rbz5 + MJD3 (control)	2	1950	209.648	171	223	10.908	0.145	127.994	0	180	29.141	0.374317971
M17 + Rbz5 + MJD3 (control)	3	1950	159.524	93	217	61.032	0.145	117.517	0	173	39.618	1.540511889
WT + Rbz8 + MJD3 (control)	4	1950	158.649	93	215	61.907	0.145	108.97	0	172	48.165	1.28531091
M13 + Rbz8 + MJD3 (control)	5	1950	202.231	149	228	18.325	0.145	106.813	0	167	50.322	0.364154843
M17 + Rbz8 + MJD3 (control)	6	1950	135.732	92	228	84.824	0.145	91.303	0	170	65.832	1.288491919
WT + Rbz8 + Rbz5	7	1950	160.446	94	220	60.11	0.145	103.588	0	163	53.547	1.122565223
M13 + Rbz8 + Rbz5	8	1950	211.355	190	228	9.201	0.145	108.827	0	165	48.308	0.190465347
M17 + Rbz8 + Rbz5	9	1950	180.063	94	221	40.493	0.145	122.097	0	166	35.038	1.15568811
Rbz8	10	1950	223.618	215	234	-3.062	0.145	94.39	0	161	62.745	-0.048800701
Rbz5	11	1950	221.66	213	229	-1.104	0.145	105.268	0	158	51.867	-0.02128521
Non-transfected	12	1950	221.986	215	245	-1.43	0.145	99.61	0	158	57.525	-0.024858757
Empty box	13	1950	220.556	215	229		0.145	157.135	142	171		

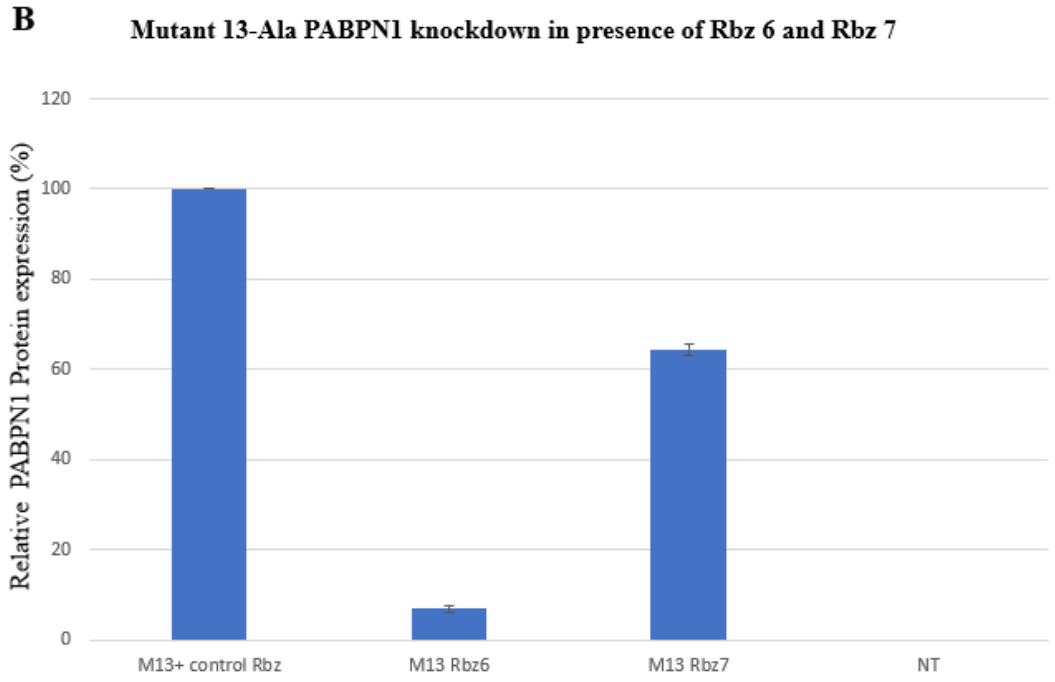
C Sample content	#	SET 1 (Empty -mean)/(Empty total protein- mean total protein)	SET 2-Duplicate (Empty - mean)/(Empty total protein- mean total protein)	SET 1 (%)	SET 2 (%)	Average of the two sets (SET 1 (%) + SET 2 (%))/2)
WT PABPN1	1	1.557770233	3.1631555	100	100	100
M13 PABPN1	2	0.344203561	1.4992948	24.462304	32.171537	28.31692034
M17 PABPN1	3	1.48366199	3.02371	95.387179	92.251119	93.81914897
WT + Rbz5	4	0.642971319	1.6528918	43.058915	48.149655	45.60428506
M13 + Rbz5	5	0.063313373	2.5273369	6.9784702	4.8687674	5.923618832
M17 + Rbz5	6	0.677148371	0.9216689	45.186245	58.174602	51.68042323
WT + MJD3 (control)	7	1.593538511	0.8271203	102.22637	95.654504	98.94043882
M13 + MJD3 (control)	8	0.360472751	0.9865621	25.474969	25.914427	25.69469809
M17 + MJD3 (control)	9	1.185014886	0.1104251	76.798077	73.669108	75.23359284
WT + Rbz8	10	0.76700098	0.7857735	50.779064	52.936149	51.85760651
M13 + Rbz8	11	0.056512899	0.2973514	6.555179	10.69389	8.624534631
M17 + Rbz8	12	0.777701959	0.1214698	51.44514	60.834488	56.13981403
WT + Rbz5 + MJD3 (control)	13	2.276893369	0.1259316	144.76137	80.186192	112.4737803
M13 + Rbz5 + MJD3 (control)	14	0.374317971	0.0316566	26.336756	5.2129502	15.7748533
M17 + Rbz5 + MJD3 (control)	15	1.540511889	2.9144965	98.925765	36.201947	67.56385618
WT + Rbz8 + MJD3 (control)	16	1.28531091	1.8209921	83.04094	30.149291	56.5951155
M13 + Rbz8 + MJD3 (control)	17	0.364154843	2.3182058	25.704159	5.3519908	15.52807478
M17 + Rbz8 + MJD3 (control)	18	1.288491919	1.906347	83.23894	33.514472	58.37670616
WT + Rbz8 + Rbz5	19	1.122565223	1.1158967	72.910937	27.202902	50.05691969
M13 + Rbz8 + Rbz5	20	0.190465347	1.0296565	14.892965	2.4141292	8.653547143
M17 + Rbz8 + Rbz5	21	1.15568811	1.1891659	74.97265	38.485212	56.72893146
Rbz8	22	-0.048800701	-0.0250881	0	0.6458145	0.322907252
Rbz5	23	-0.02128521	-0.0458121	1.7126845	0	0.856342238
Non-transfected	24	-0.024858757	-0.0419353	1.4902513	0.1208116	0.805531459

Figure 11. Sample western blot calculations shown for Rbz5 and Rbz8 samples

To quantify the protein present in cells in each sample, the intensity of the pixels of each band on the protein blots in a defined area was measured. Two identical samples were run on gel and two identical sets of blots were prepared for quantification. Table A shows the measurements for the first and table B shows the same values for set 2. ImageJ software measures the maximum, minimum, and the mean intensity of each band as shown in tables A and B. The area of these measurements on each blot for all bands are kept constant shown in area columns in tables A and B. The mean intensity values were used for all calculations. The pixel intensity of the background, labeled as empty box was subtracted from the measured mean intensities in A and B. Next, the pixel intensity of each of the total protein content columns associated with the bands (same area for all columns on one plot) was measured and the pixel intensity of the background of the corresponding blot was subtracted from these values. The mean intensity of each band was then divided by the mean intensity of the corresponding total protein column to ensure the band intensities are normalized to the total protein loaded on the gel and are comparable.
In table C the values obtained from each duplicated blot (SET 2 calculations are not shown) was normalized to a percentage. In the last step of calculations, the average of the values obtained from the duplicated blots was calculated and used to generate the relative protein expression graphs.

A3. Protein expression in presence of Ribozymes 6 and 7





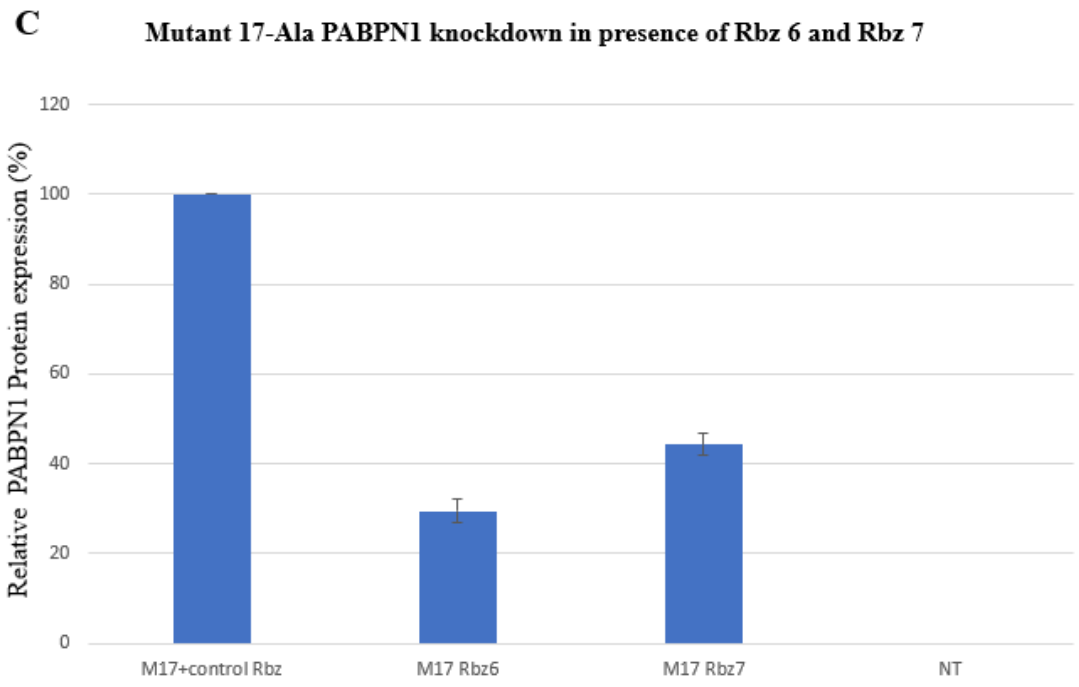
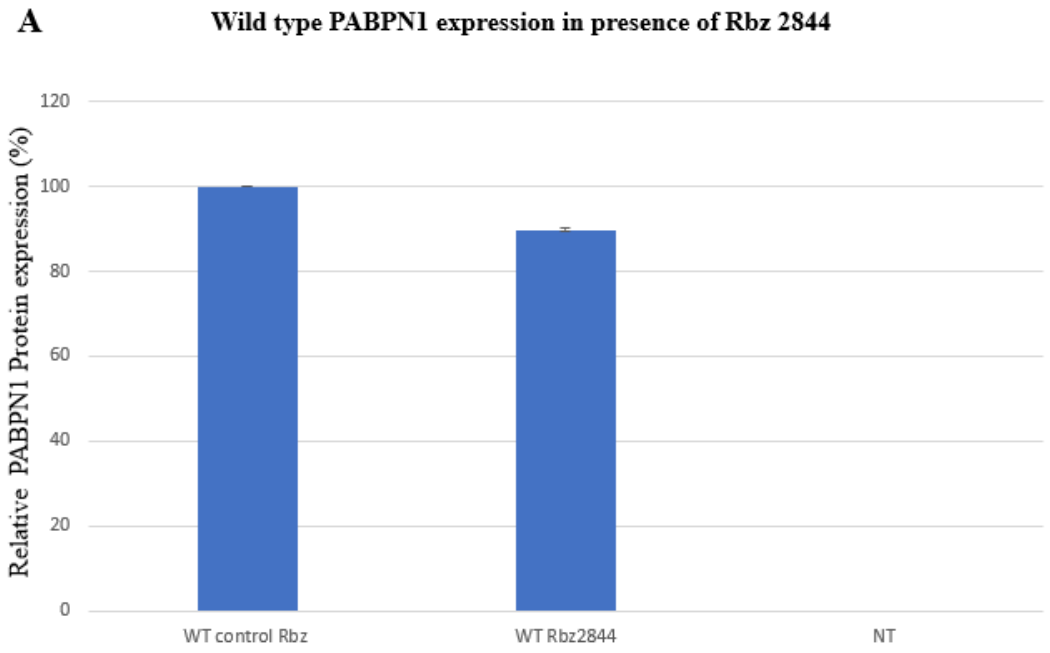
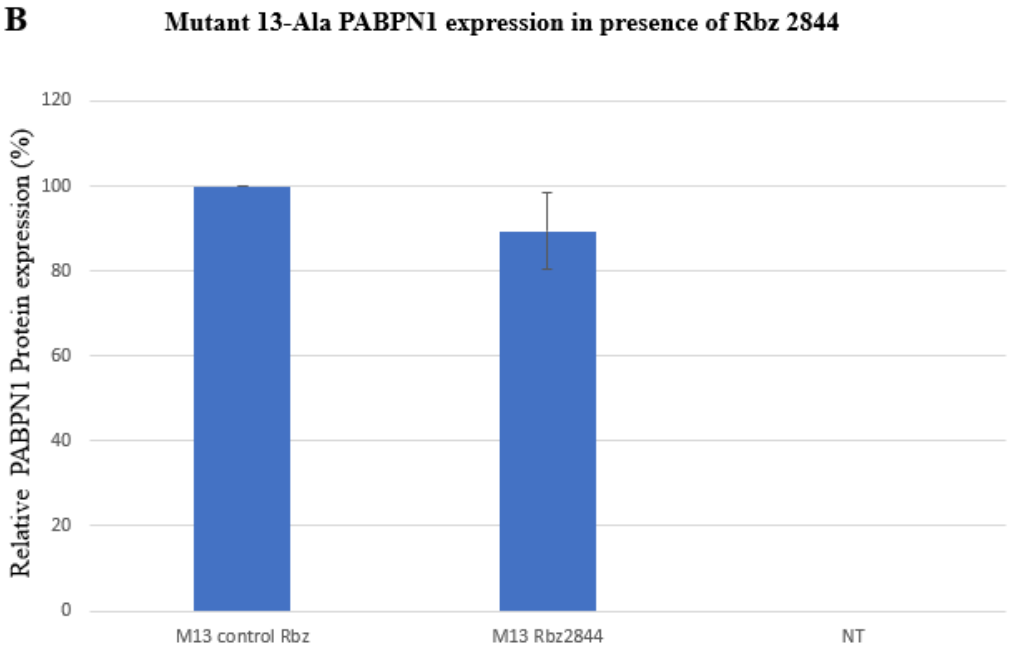


Figure 12. Non-selective PABPN1 protein knockdown in presence in ribozyme 6 and ribozyme 7

M13 represents the mutant PABPN1 with a 13 alanine repeat while M17 is the mutant with 17 alanine residues and WT is the wild type gene with a 10 Alanie repeat. The control ribozyme does not target PABPN1 gene. As depicted in the graphs, the knockdown of the proteins when cells were co-transfected with Rbz6 and Rbz7 is indiscriminate, and the ribozyme appears to cleave wild type protein as well as the two mutant types. It is important to note that this group of ribozymes are efficient in binding to the targets and capable of cleaving the transcript but not selective as they can also bind and cleave the wild type.

A4. Protein expression in presence of ribozyme 2844





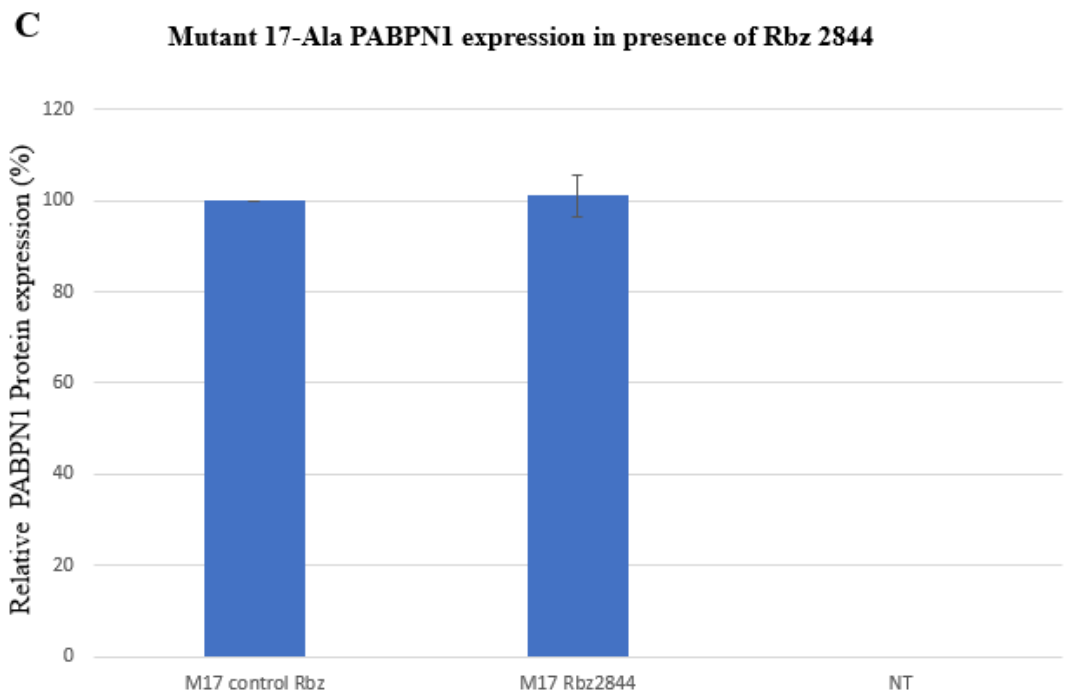
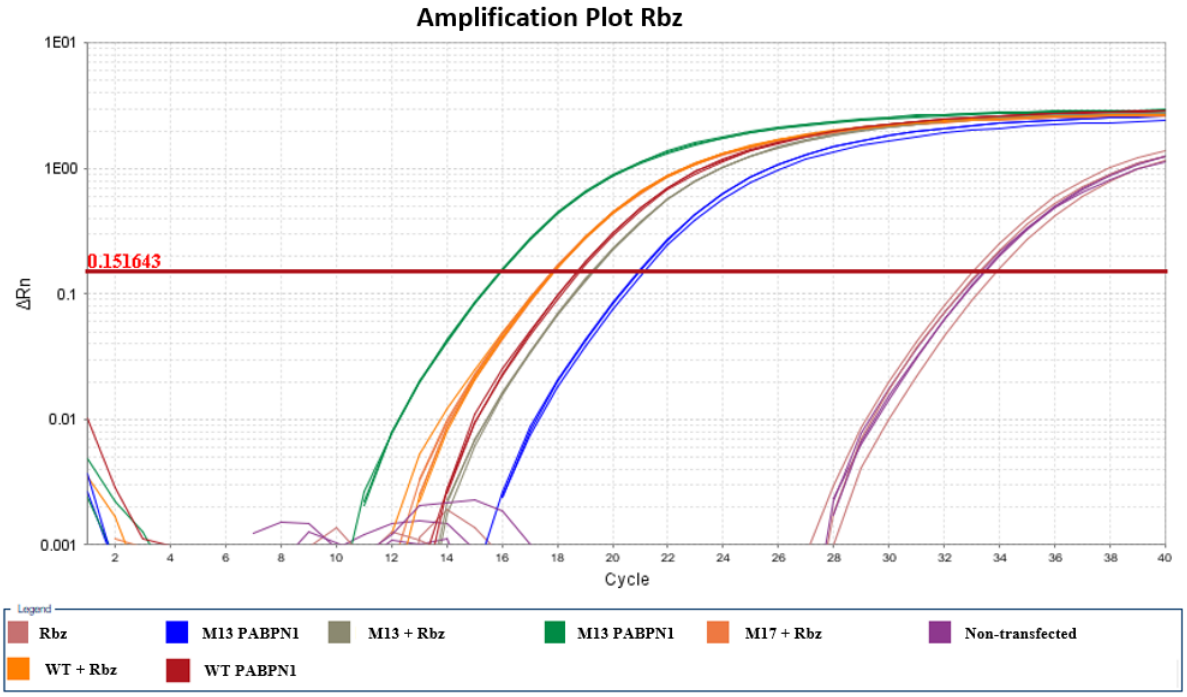
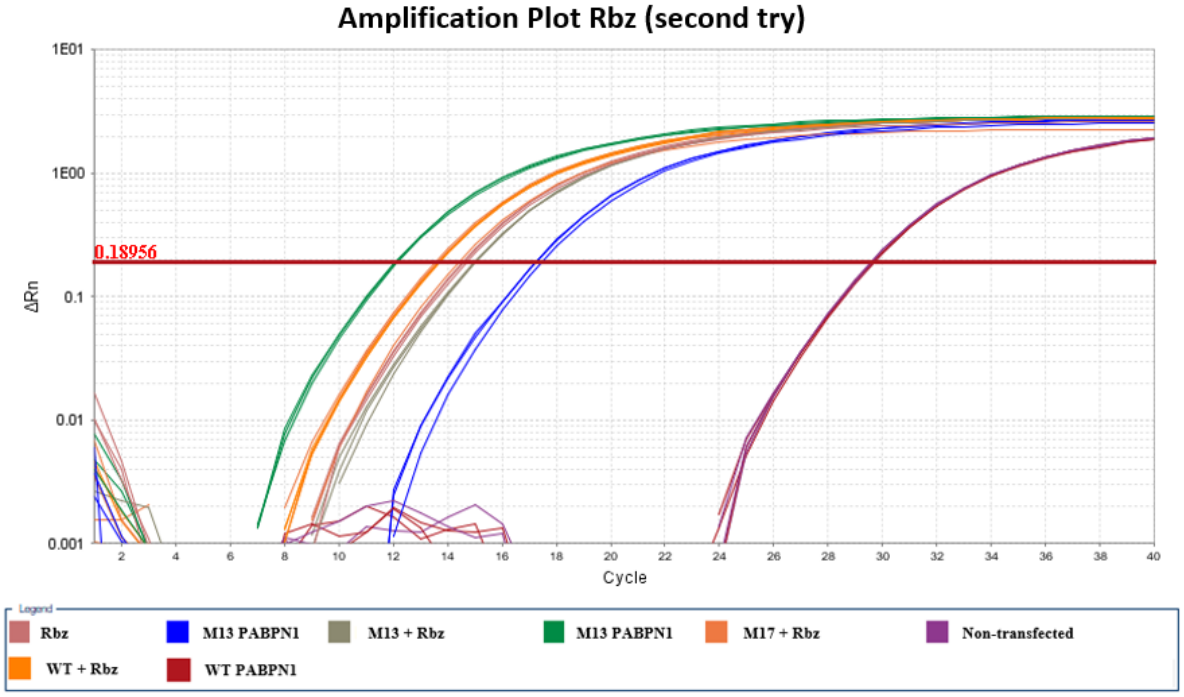


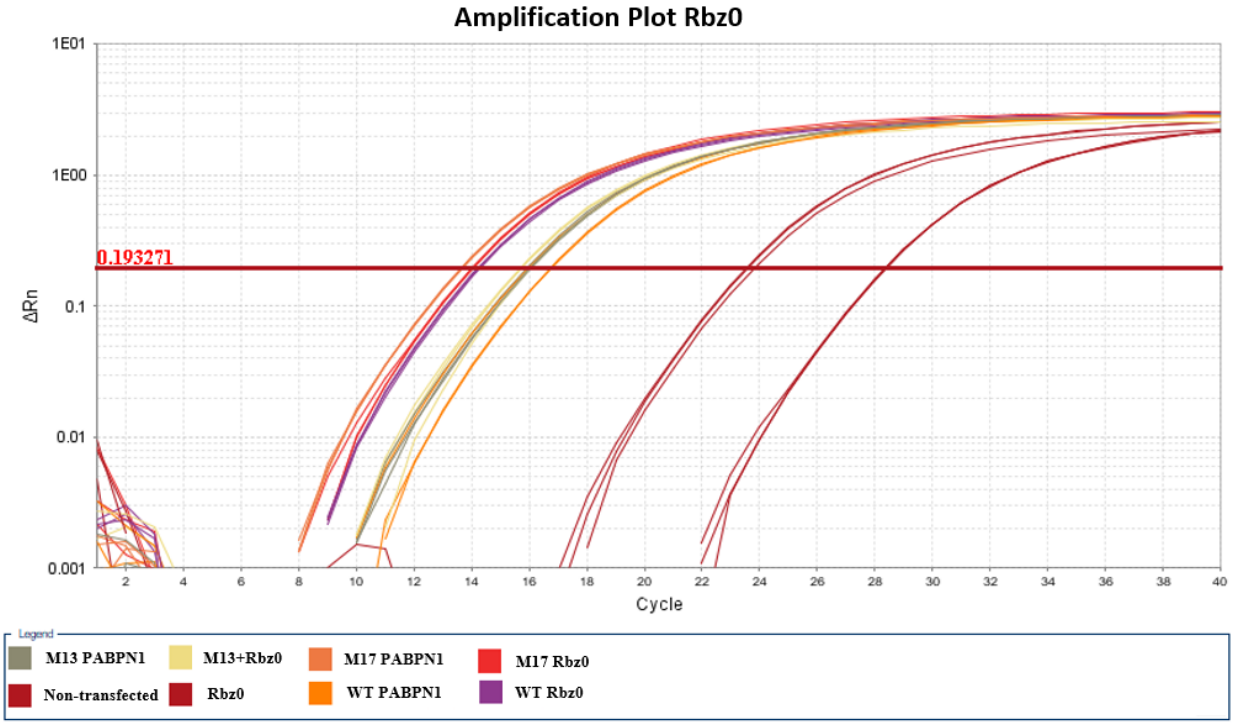
Figure 13. PABPN1 protein expression in presence in ribozyme 2844

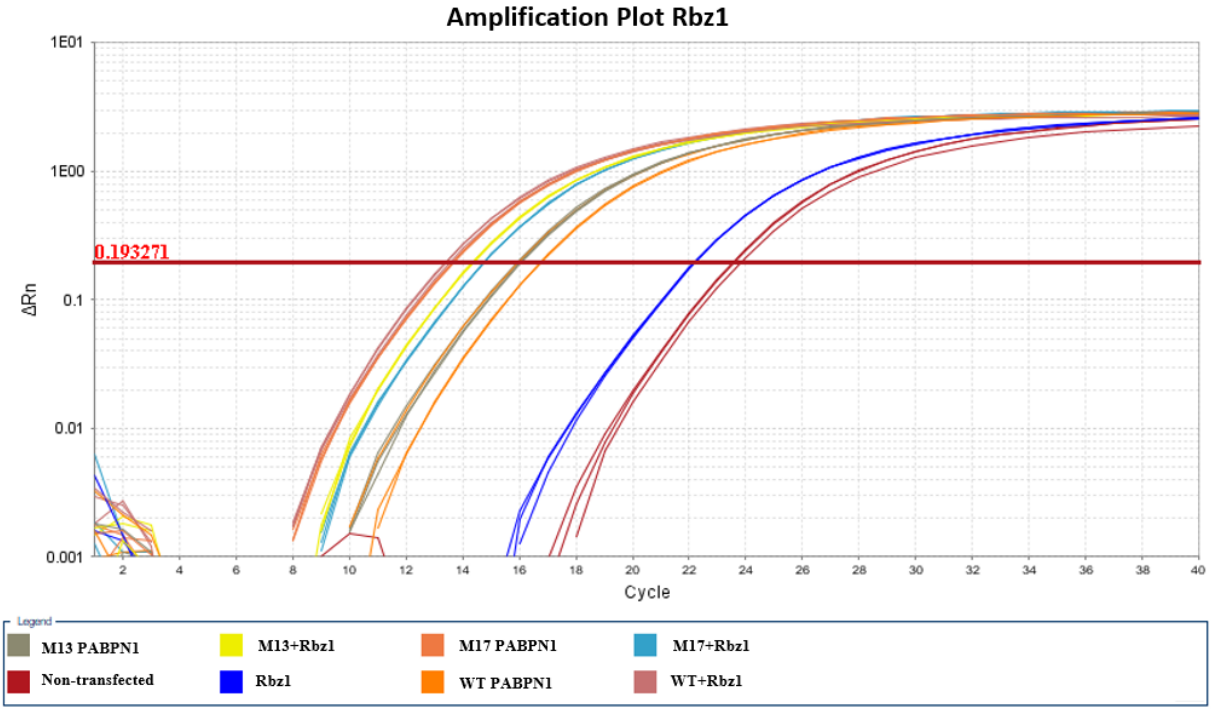
M13 represents the mutant PABPN1 with a 13 alanine repeat while M17 is the mutant with 17 alanine residues and WT is the wild type gene with a 10-alanine repeat. The control ribozyme does not target PABPN1 gene. Rbz2844 does not seem to cleave any of the transcripts of PABPN1 to an observable degree and is hence categorized as inefficient.

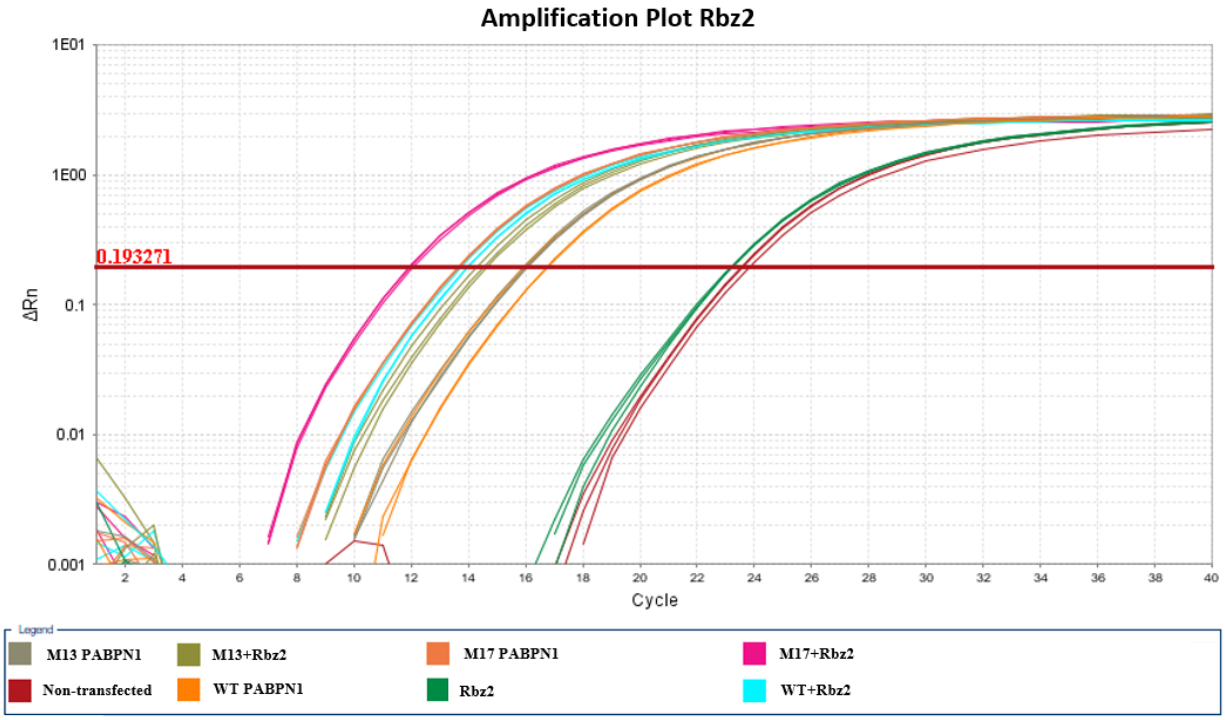
A5. RT-qPCR amplification plots:

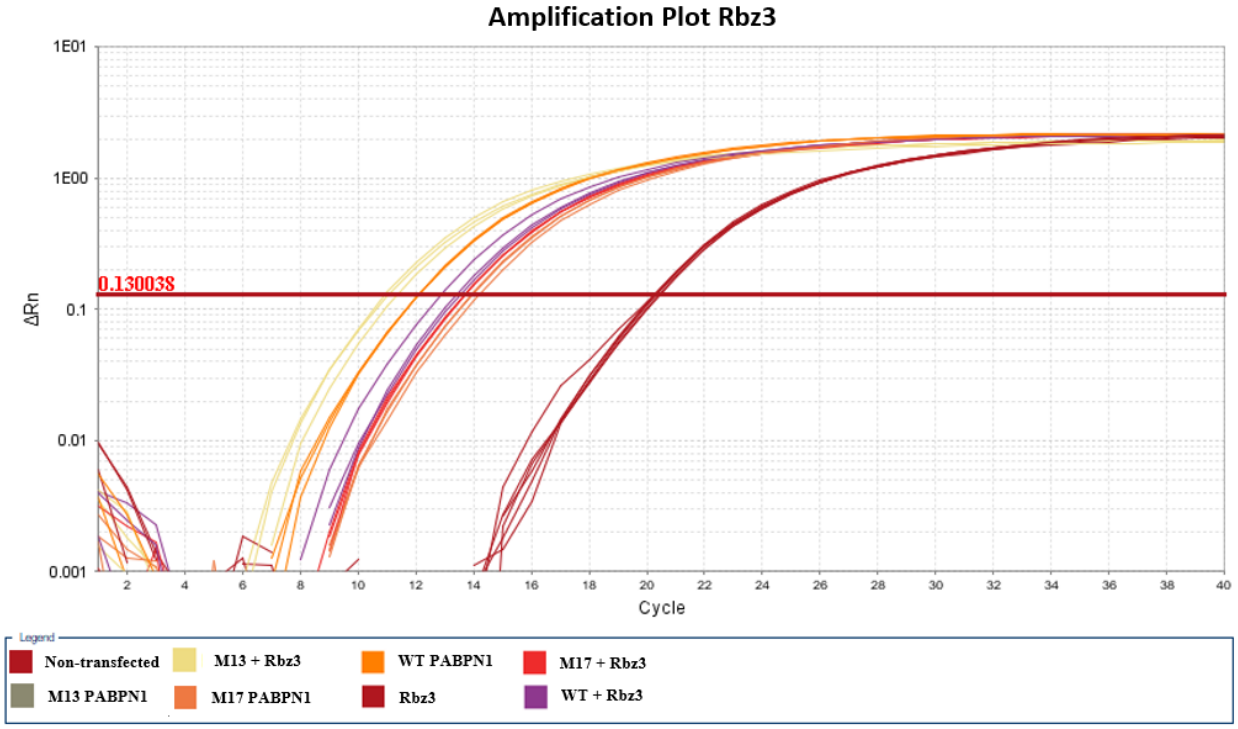


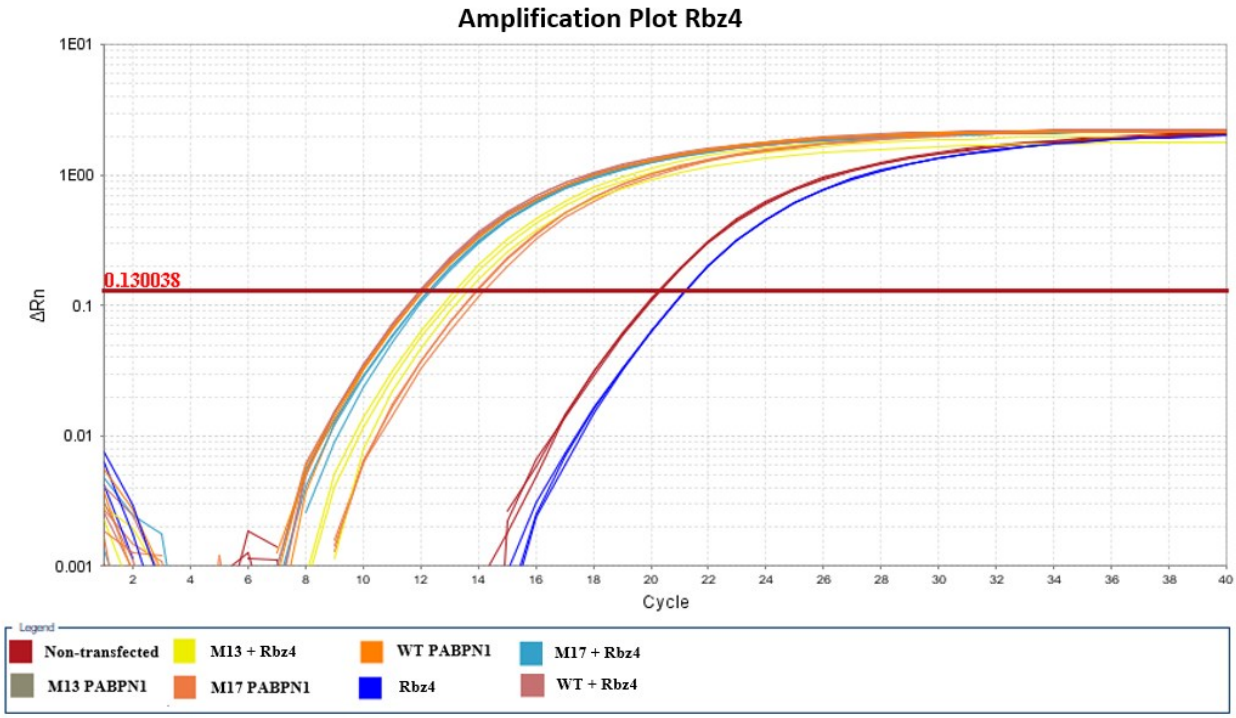


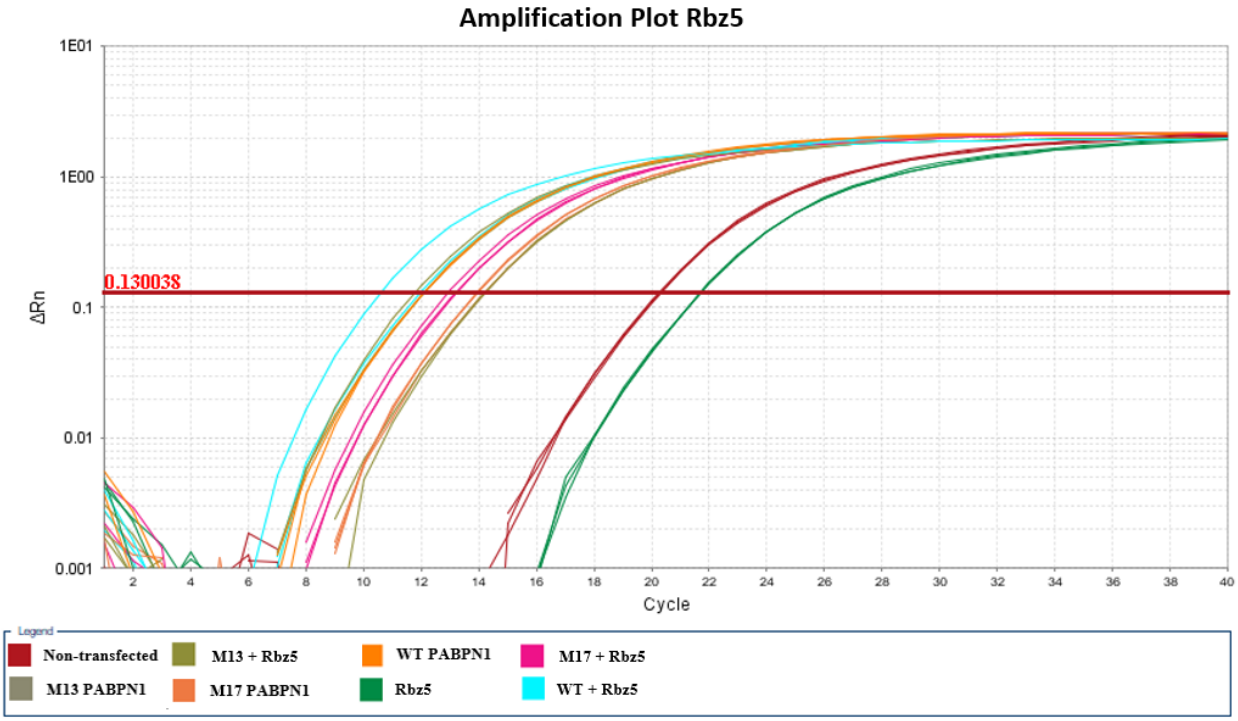

Amplification Plot Rbz0
1E01
1E00
0.193271
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+Rbz0
M17 PABPN1
M17 Rbz0
Non-transfected
Rbz0
WT PABPN1
WT Rbz0

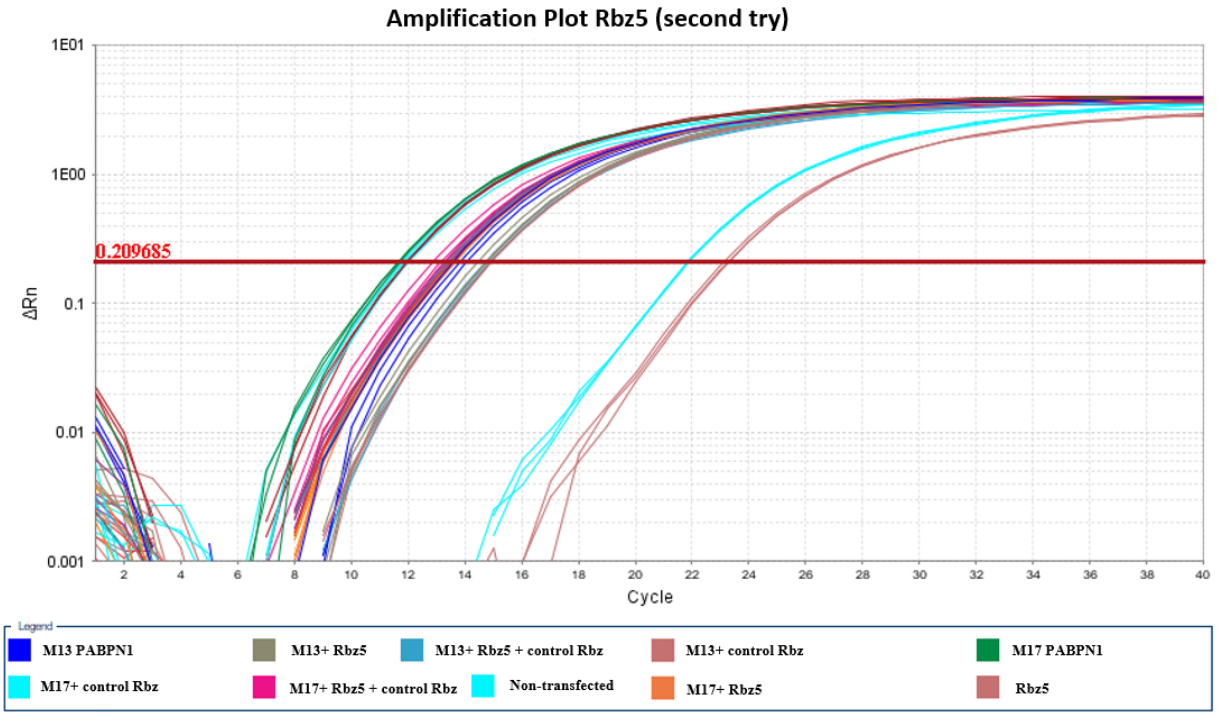

Amplification Plot Rbz1
1E01
1E00
0.193271
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+Rbz1
M17 PABPN1
M17+Rbz1
Non-transfected
Rbz1
WT PABPN1
WT+Rbz1

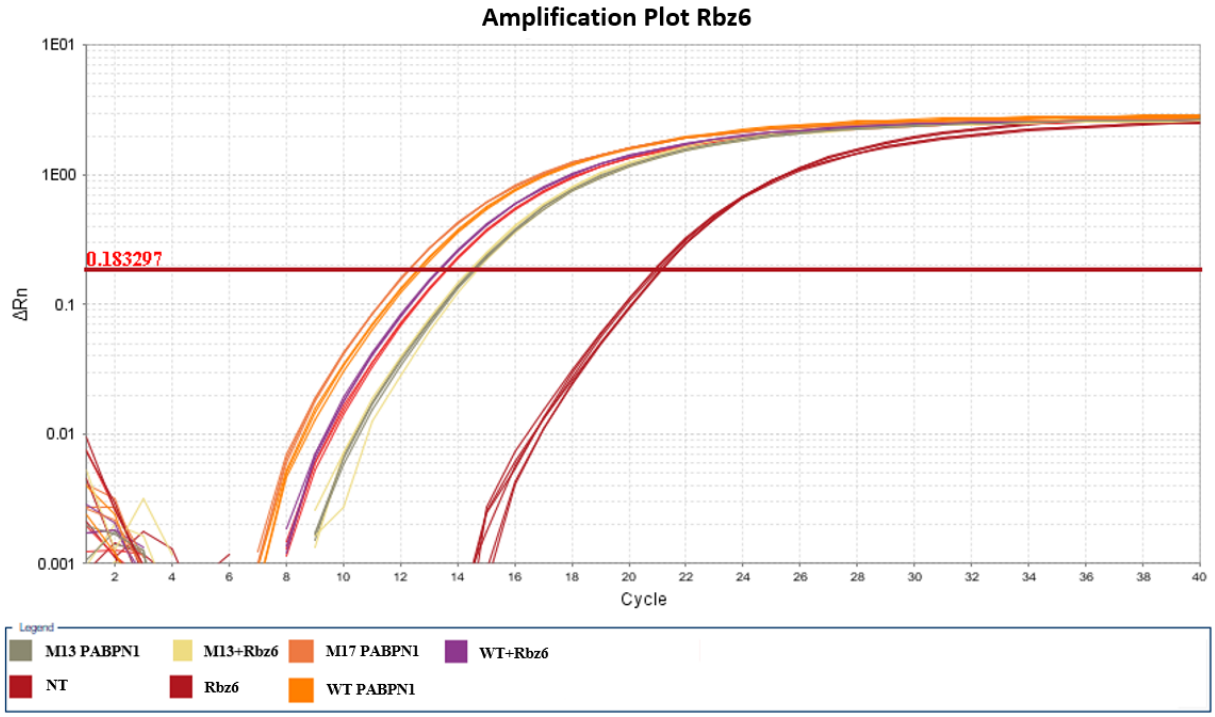

Amplification Plot Rbz2
1E01
1E00
0.193271
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+Rbz2
M17 PABPN1
M17+Rbz2
Non-transfected
WT PABPN1
Rbz2
WT+Rbz2

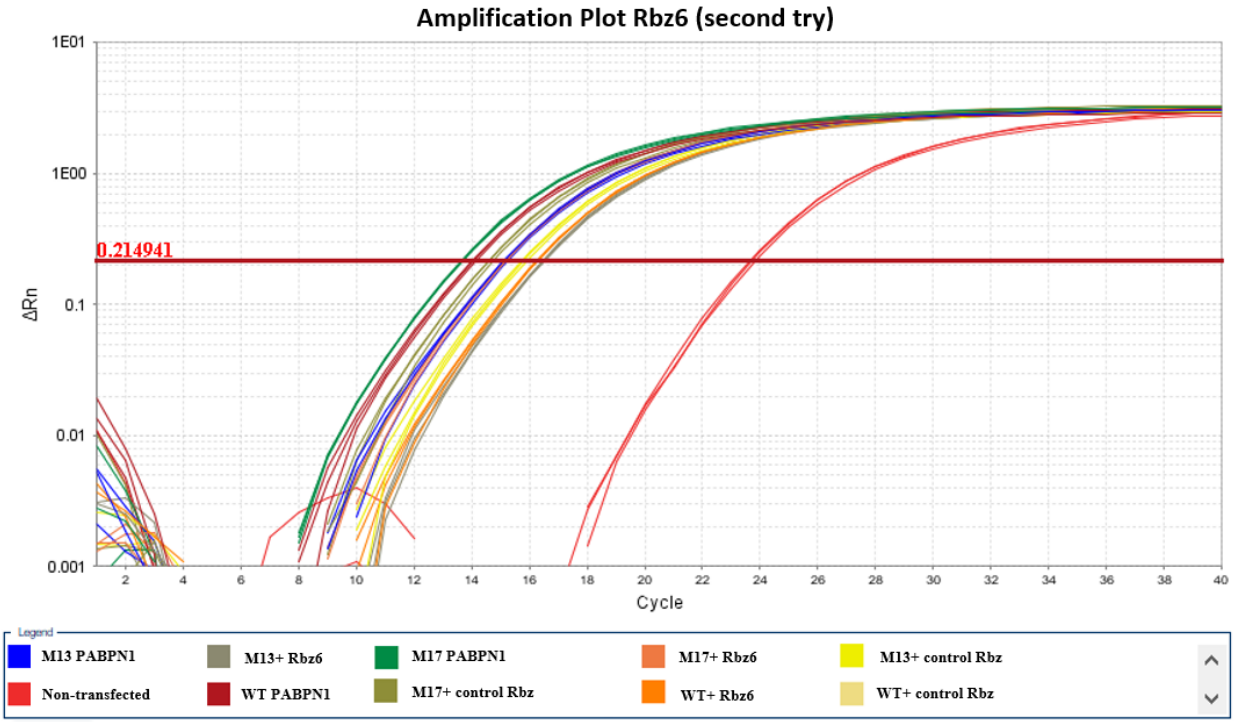

Amplification Plot Rbz3
1E01
1E00
0.130038
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
Non-transfected
M13 + Rbz3
WT PABPN1
M17 + Rbz3
M13 PABPN1
M17 PABPN1
Rbz3
WT + Rbz3

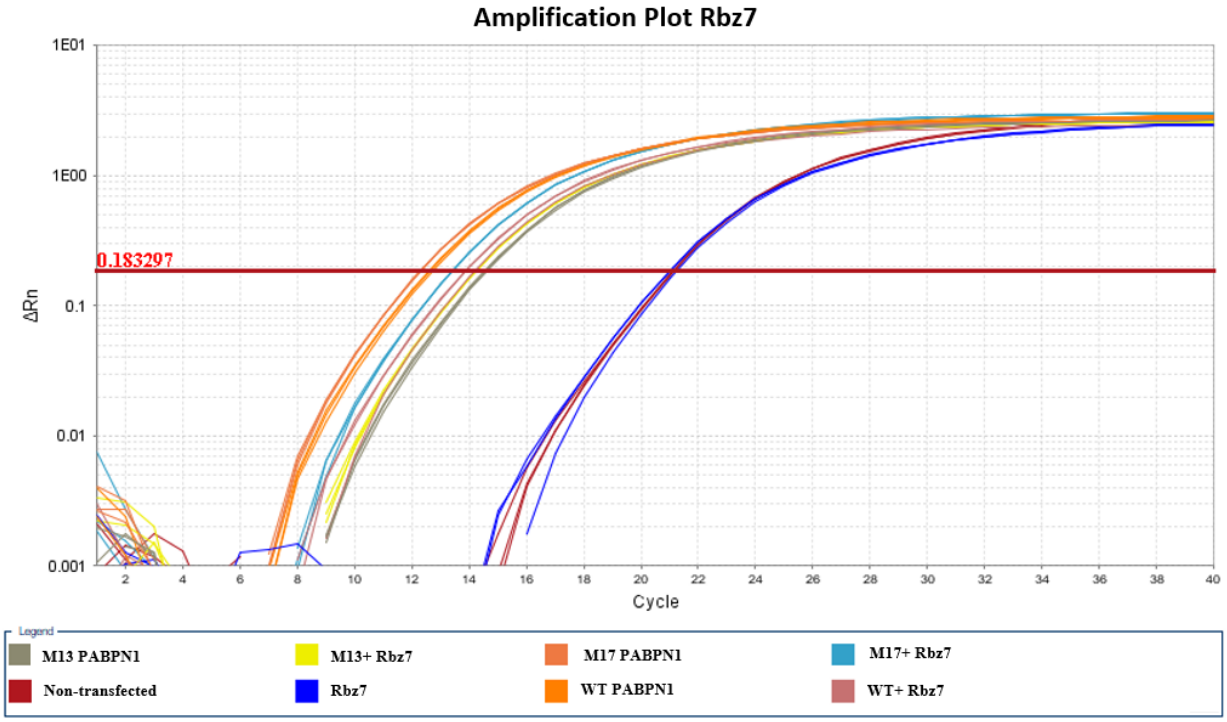

Amplification Plot Rbz4
1E01
1E00
0.130038
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
Non-transfected
M13 + Rbz4
WT PABPN1
M17 + Rbz4
M13 PABPN1
M17 PABPN1
Rbz4
WT + Rbz4

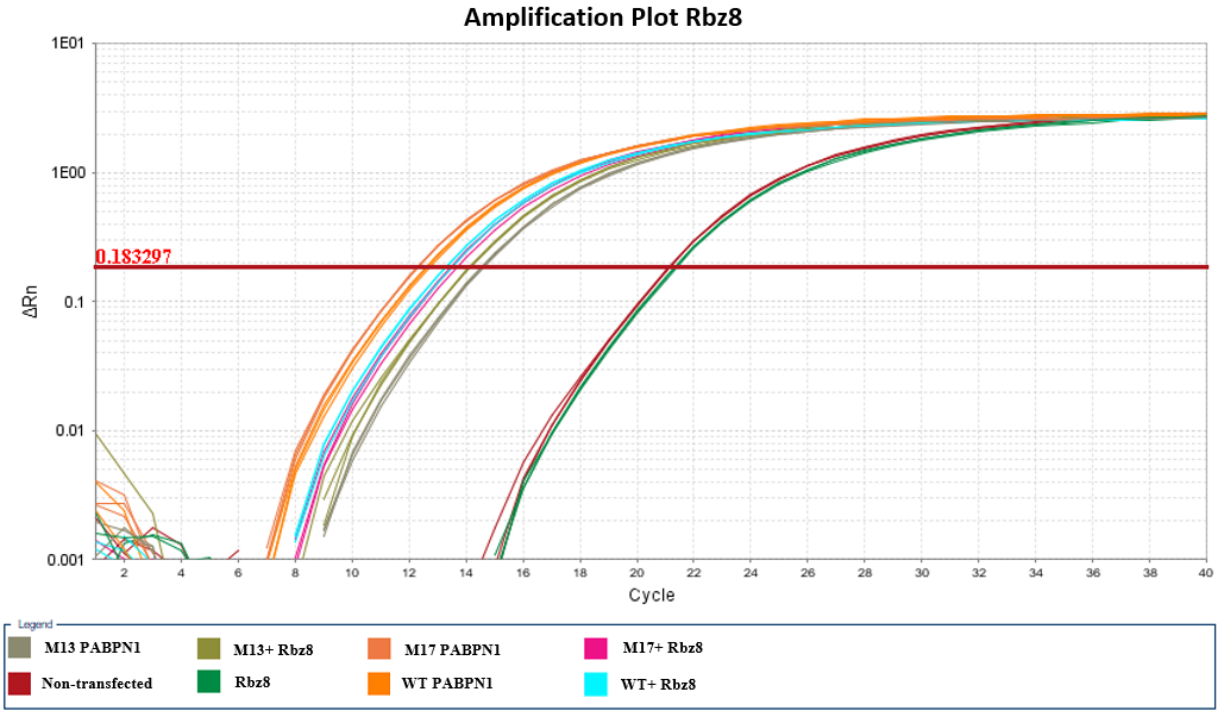

Amplification Plot Rbz5
1E01
1E00
0.130038
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
Non-transfected
M13 + Rbz5
WT PABPN1
M17 + Rbz5
M13 PABPN1
M17 PABPN1
Rbz5
WT + Rbz5

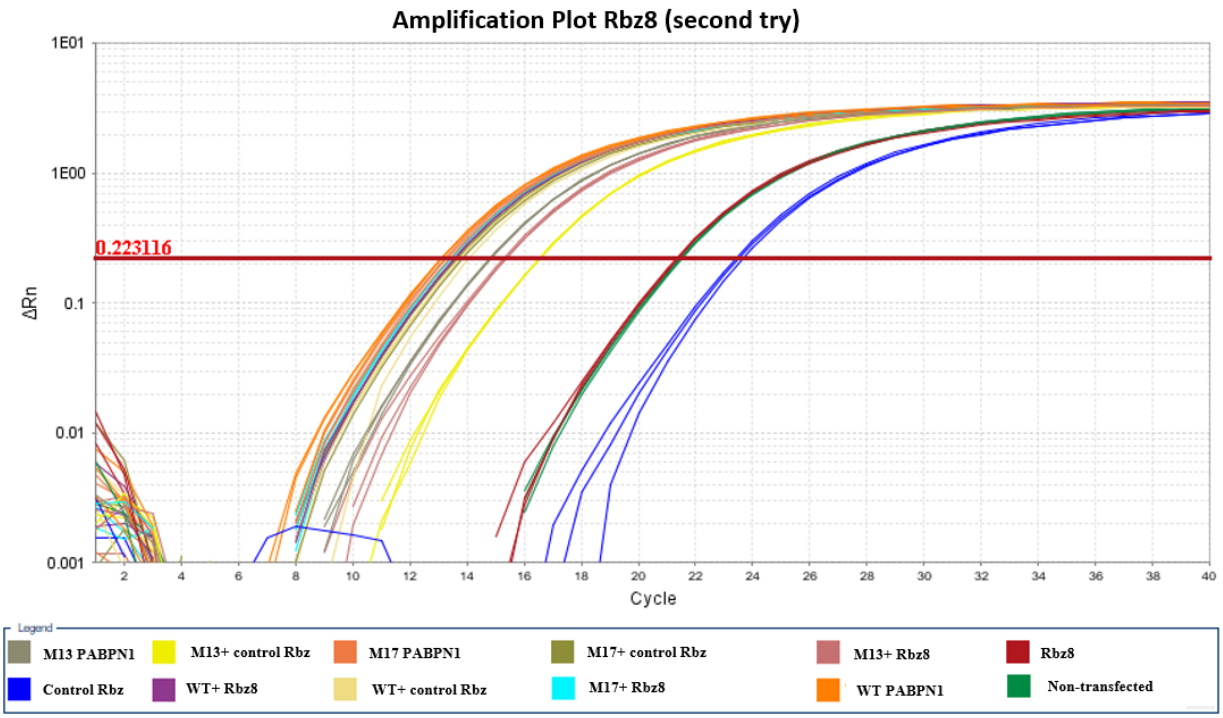

Amplification Plot Rbz5 (second try)
1E01
1E00
0.209685
0.1
0.01
0.001
ΔRn
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
Cycle
Legend
M13 PABPN1
M13+ Rbz5
M13+ Rbz5 + control Rbz
M13+ control Rbz
M17 PABPN1
M17+ control Rbz
M17+ Rbz5 + control Rbz
Non-transfected
M17+ Rbz5
Rbz5

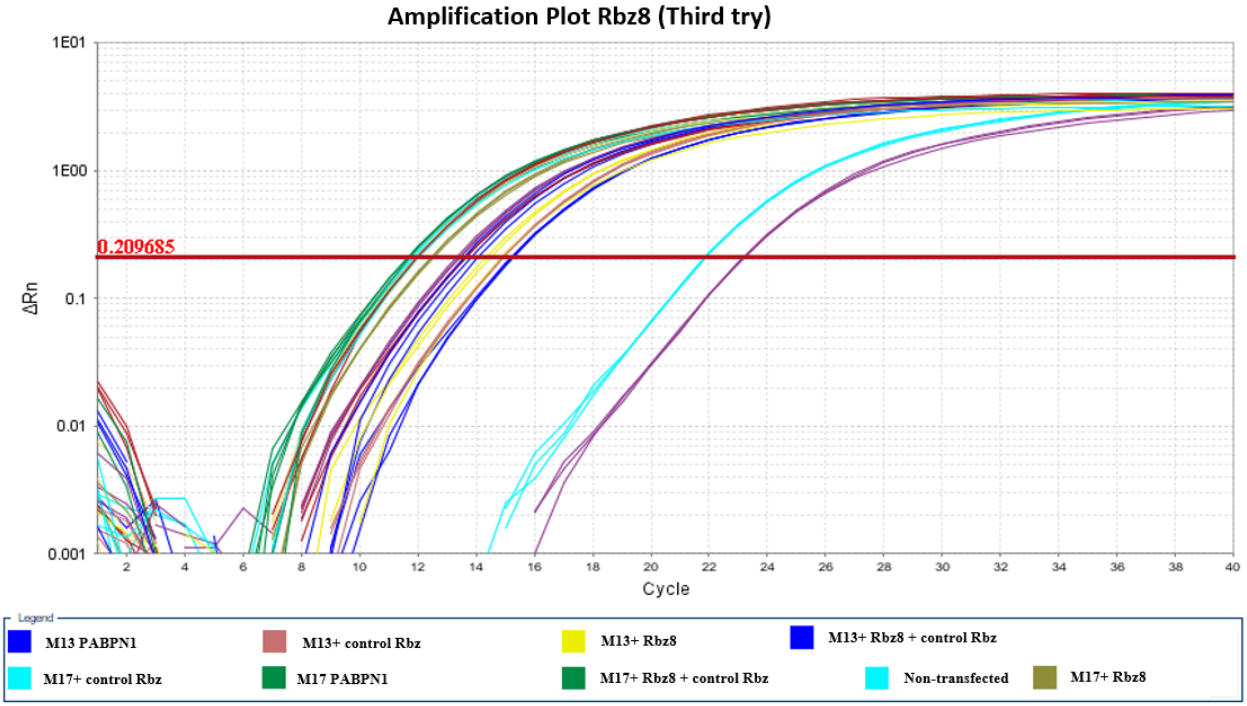

Amplification Plot Rbz6
1E01
1E00
0.183297
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+Rbz6
M17 PABPN1
WT+Rbz6
NT
Rbz6
WT PABPN1

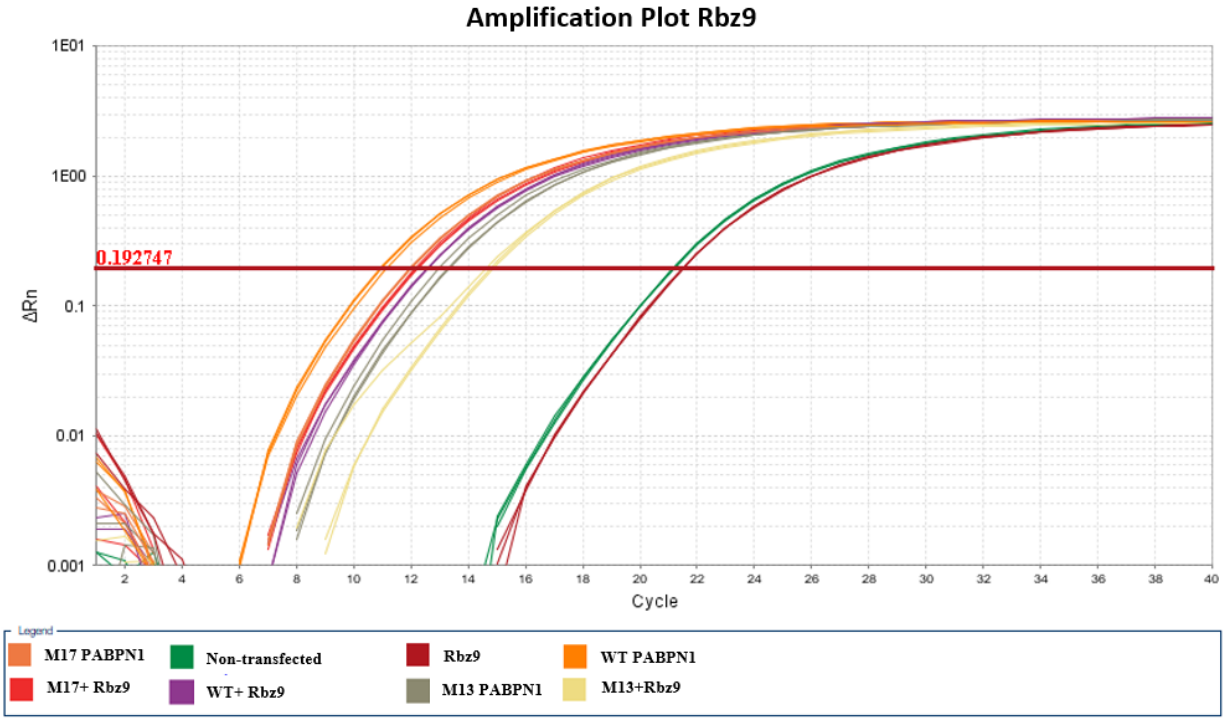

Amplification Plot Rbz6 (second try)
1E01
1E00
0.214941
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz6
M17 PABPN1
M17+ Rbz6
M13+ control Rbz
Non-transfected
WT PABPN1
M17+ control Rbz
WT+ Rbz6
WT+ control Rbz

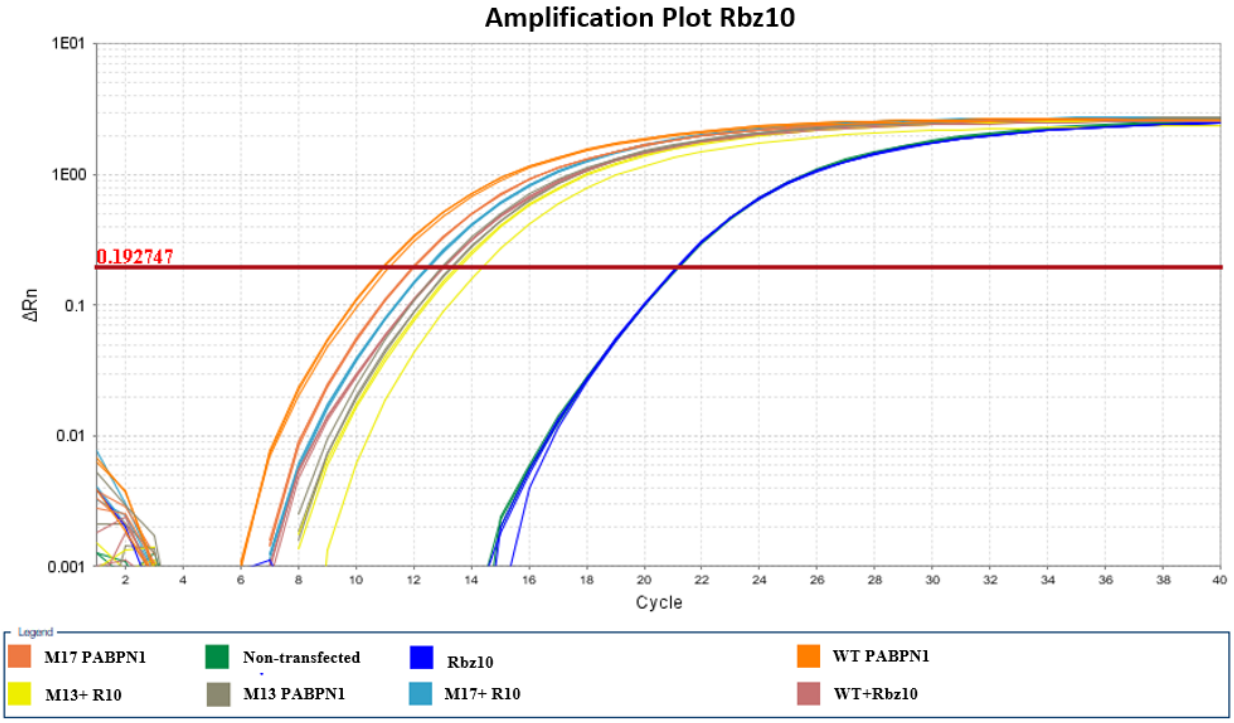

Amplification Plot Rbz7
1E01
1E00
0.183297
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz7
M17 PABPN1
M17+ Rbz7
Non-transfected
Rbz7
WT PABPN1
WT+ Rbz7

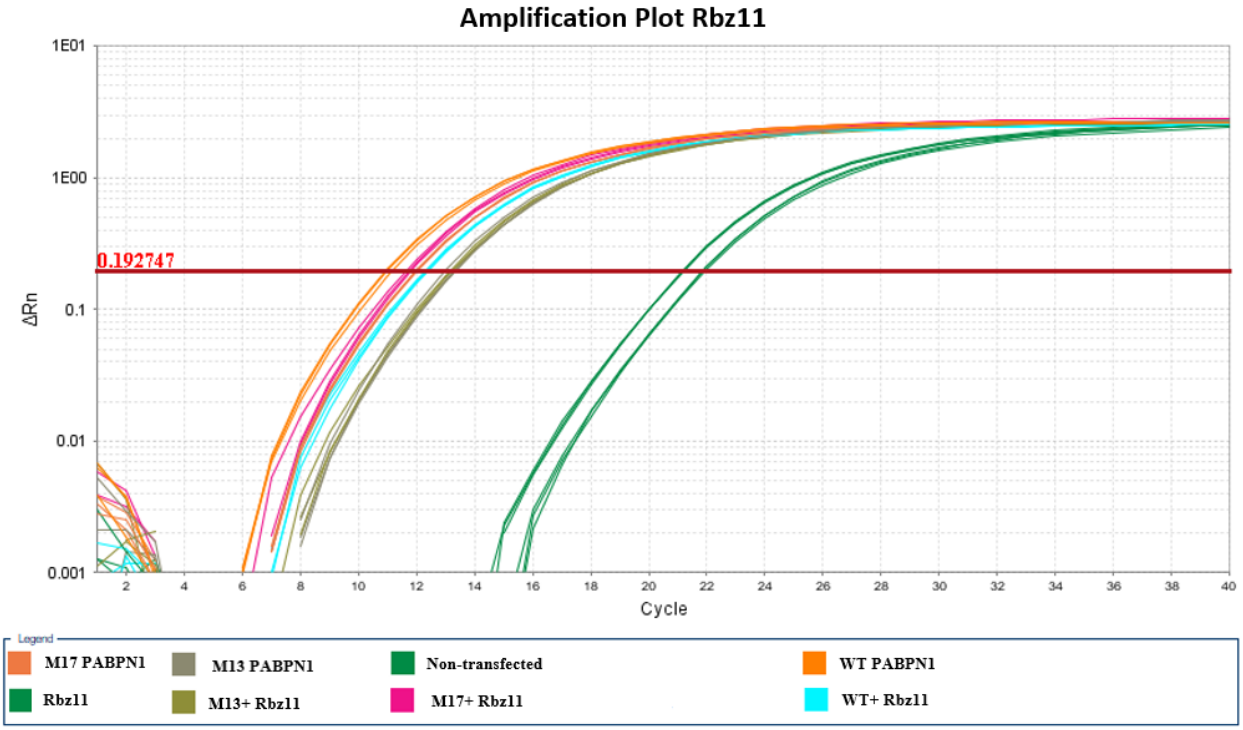

Amplification Plot Rbz8
1E01
1E00
0.183297
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz8
M17 PABPN1
M17+ Rbz8
Non-transfected
Rbz8
WT PABPN1
WT+ Rbz8

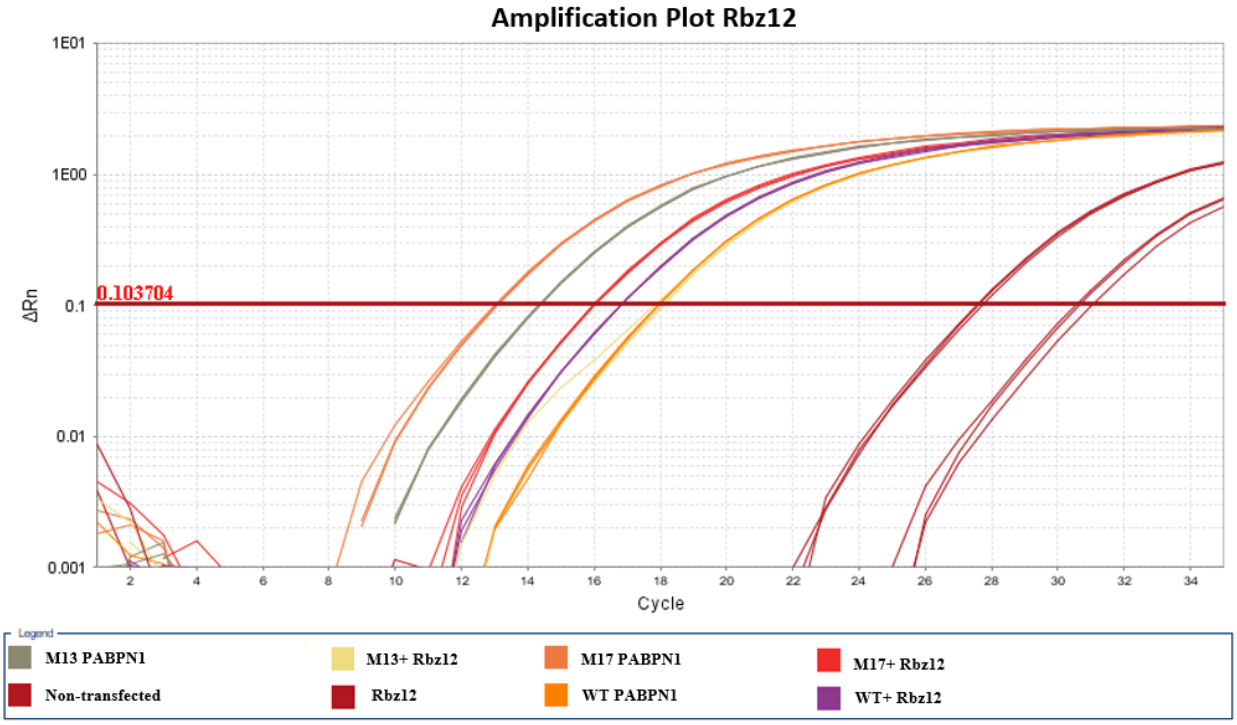

Amplification Plot Rbz8 (second try)
1E01
1E00
0.223116
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ control Rbz
M17 PABPN1
M17+ control Rbz
M13+ Rbz8
Rbz8
Control Rbz
WT+ Rbz8
WT+ control Rbz
M17+ Rbz8
WT PABPN1
Non-transfected

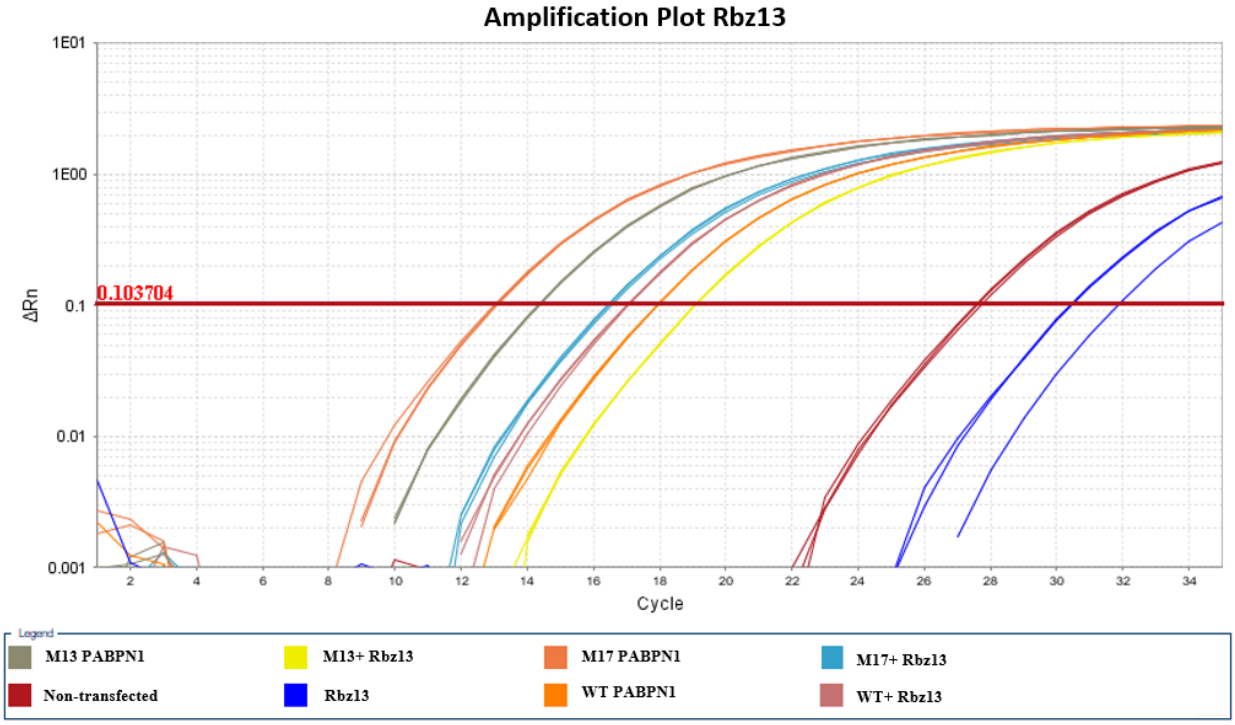

Amplification Plot Rbz8 (Third try)
1E01
1E00
0.209685
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ control Rbz
M13+ Rbz8
M13+ Rbz8 + control Rbz
M17+ control Rbz
M17 PABPN1
M17+ Rbz8 + control Rbz
Non-transfected
M17+ Rbz8

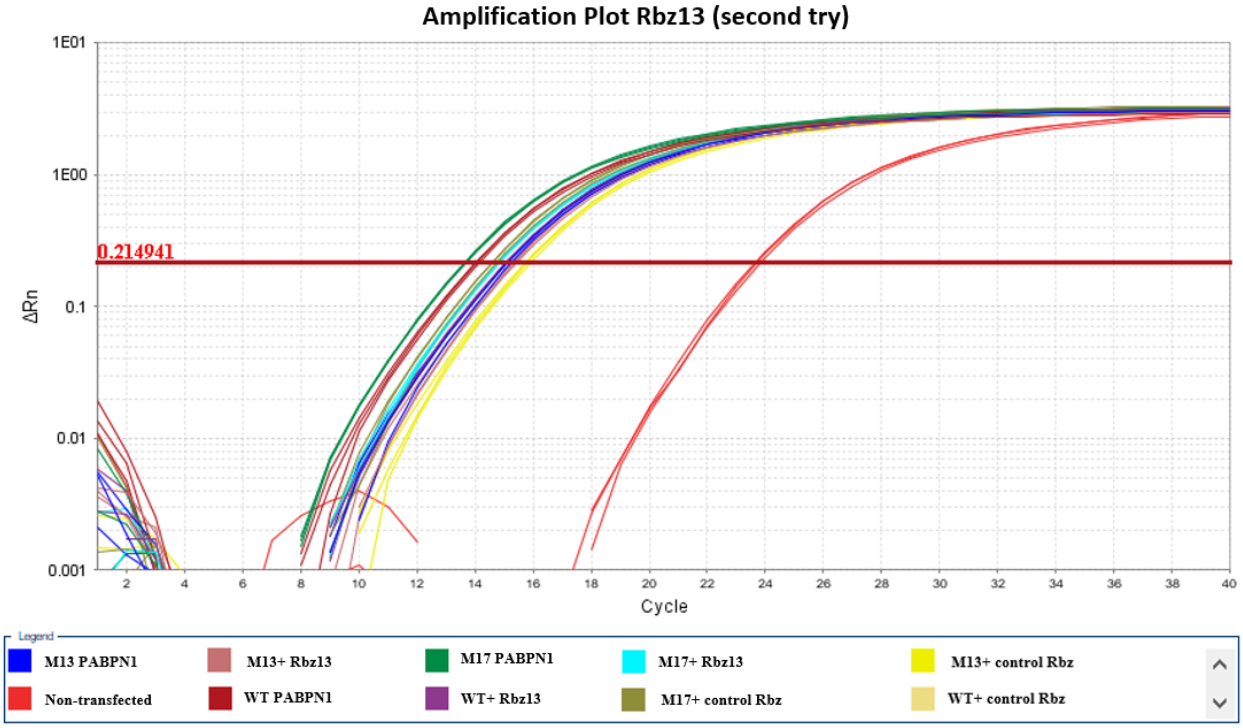

Amplification Plot Rbz9
1E01
1E00
0.192747
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
Non-transfected
Rbz9
WT PABPN1
M17+ Rbz9
WT+ Rbz9
M13 PABPN1
M13+Rbz9

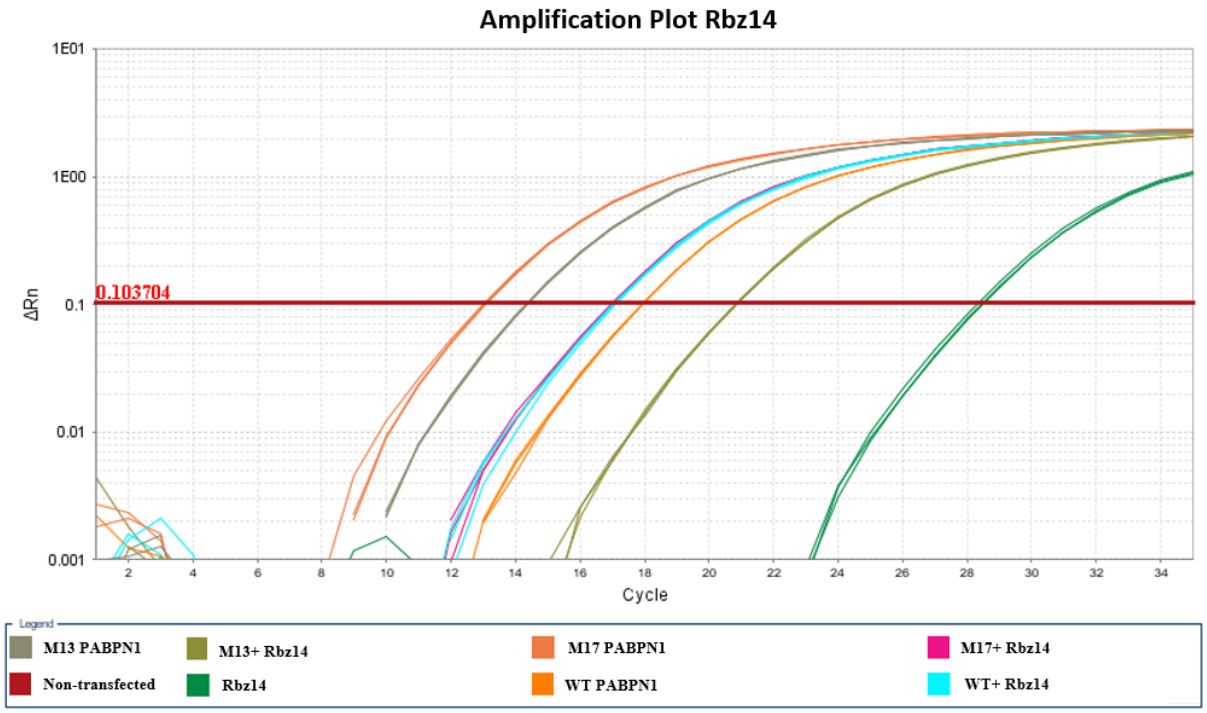

Amplification Plot Rbz10
1E01
1E00
0.192747
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
Non-transfected
Rbz10
WT PABPN1
M13+ R10
M13 PABPN1
M17+ R10
WT+Rbz10

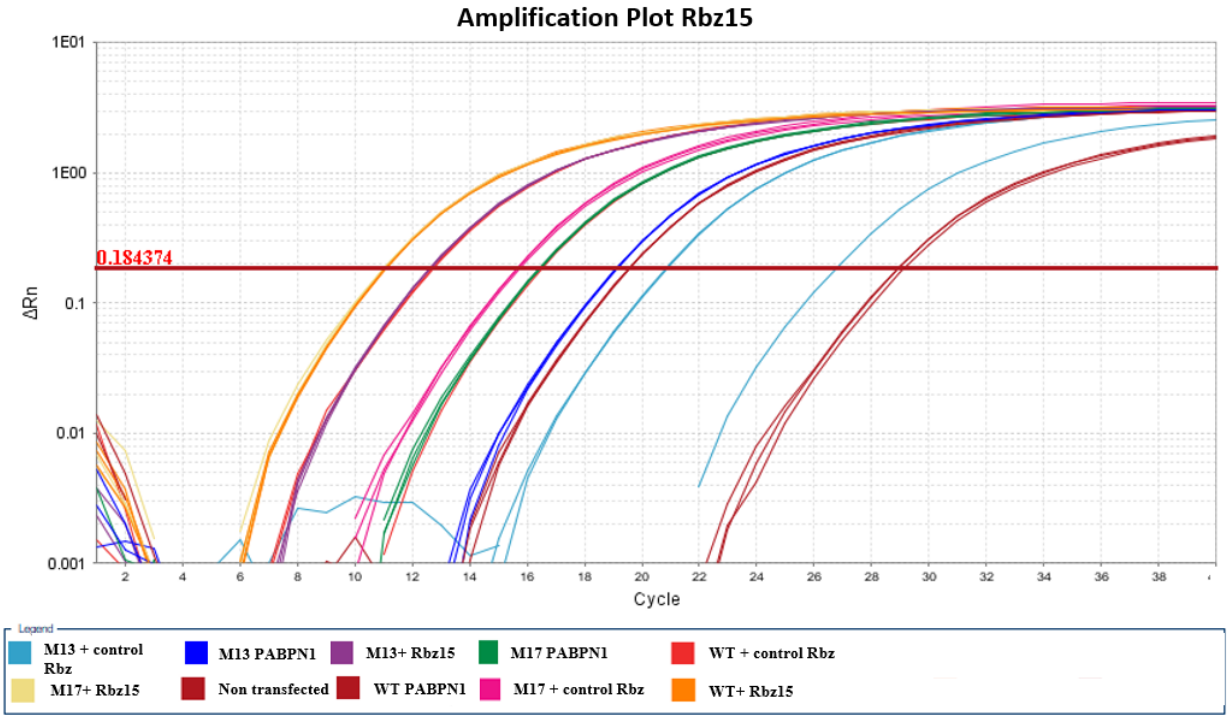

Amplification Plot Rbz11
1E01
1E00
0.192747
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
M13 PABPN1
Non-transfected
WT PABPN1
Rbz11
M13+ Rbz11
M17+ Rbz11
WT+ Rbz11

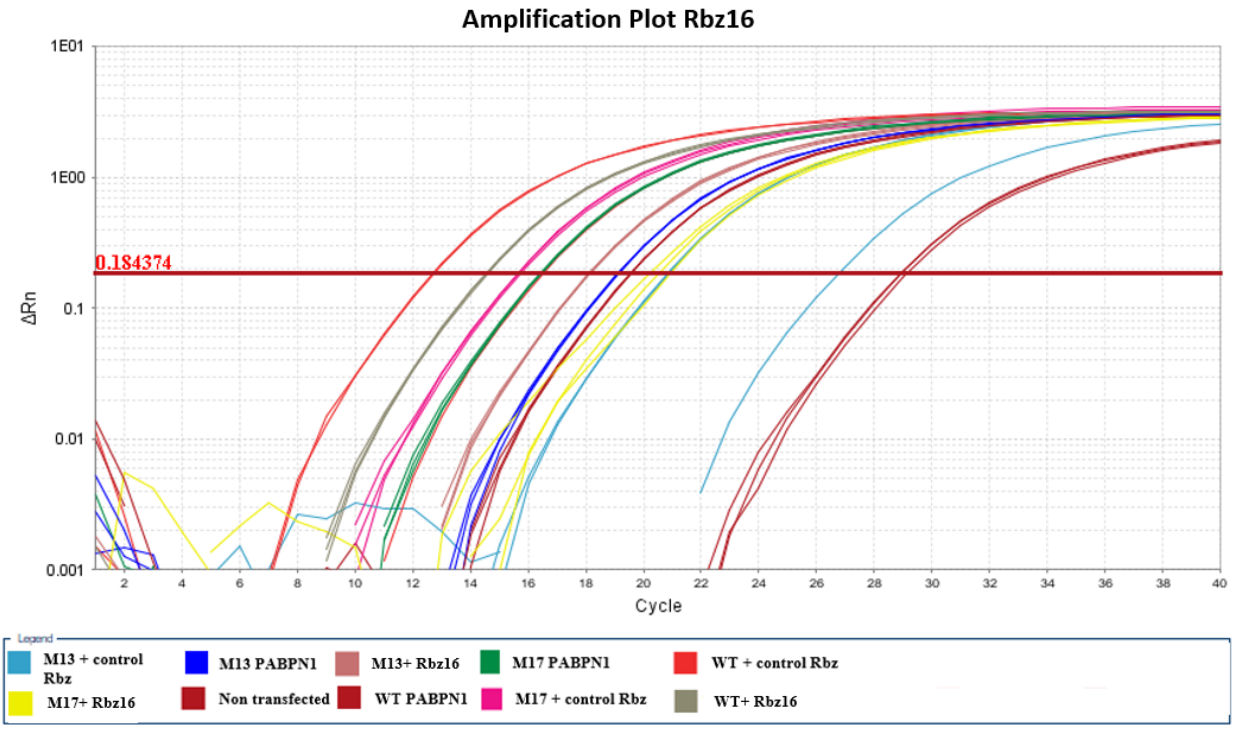

Amplification Plot Rbz12
1E01
1E00
0.103704
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
Cycle
Legend
M13 PABPN1
M13+ Rbz12
M17 PABPN1
M17+ Rbz12
Non-transfected
Rbz12
WT PABPN1
WT+ Rbz12

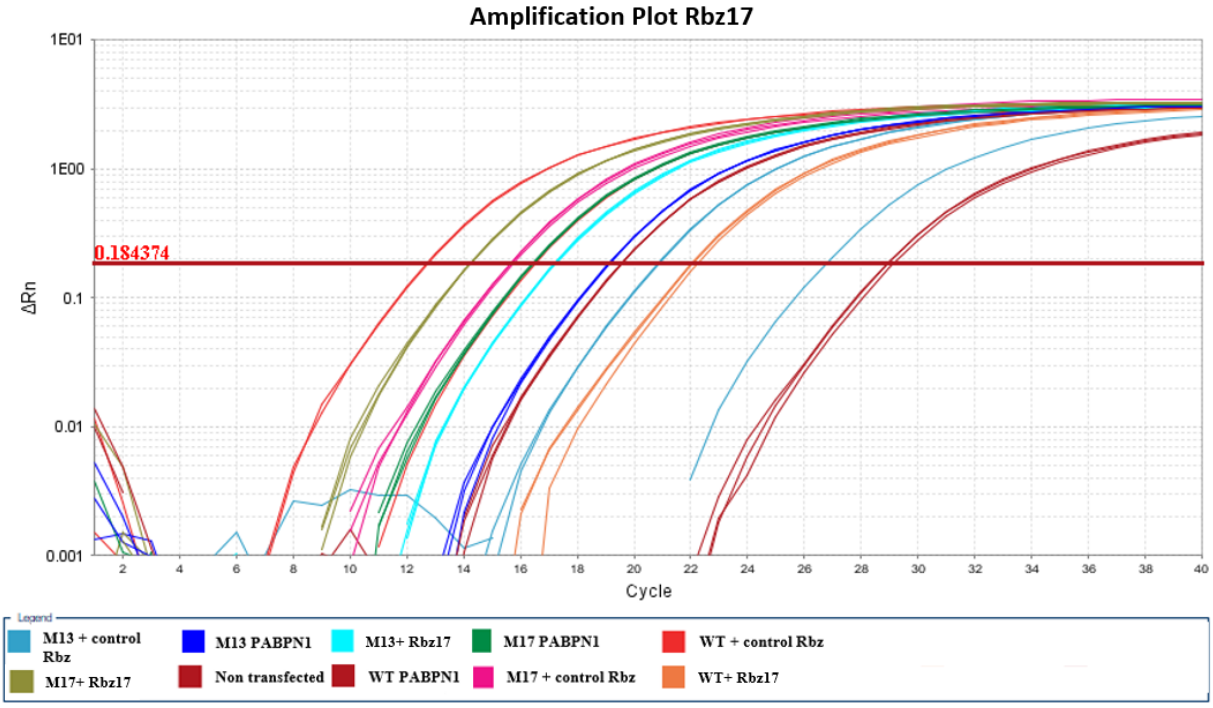

Amplification Plot Rbz13
1E01
1E00
0.1
0.01
0.001
ΔRn
0.103704
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
Cycle
Legend
M13 PABPN1
M13+ Rbz13
M17 PABPN1
M17+ Rbz13
Non-transfected
Rbz13
WT PABPN1
WT+ Rbz13

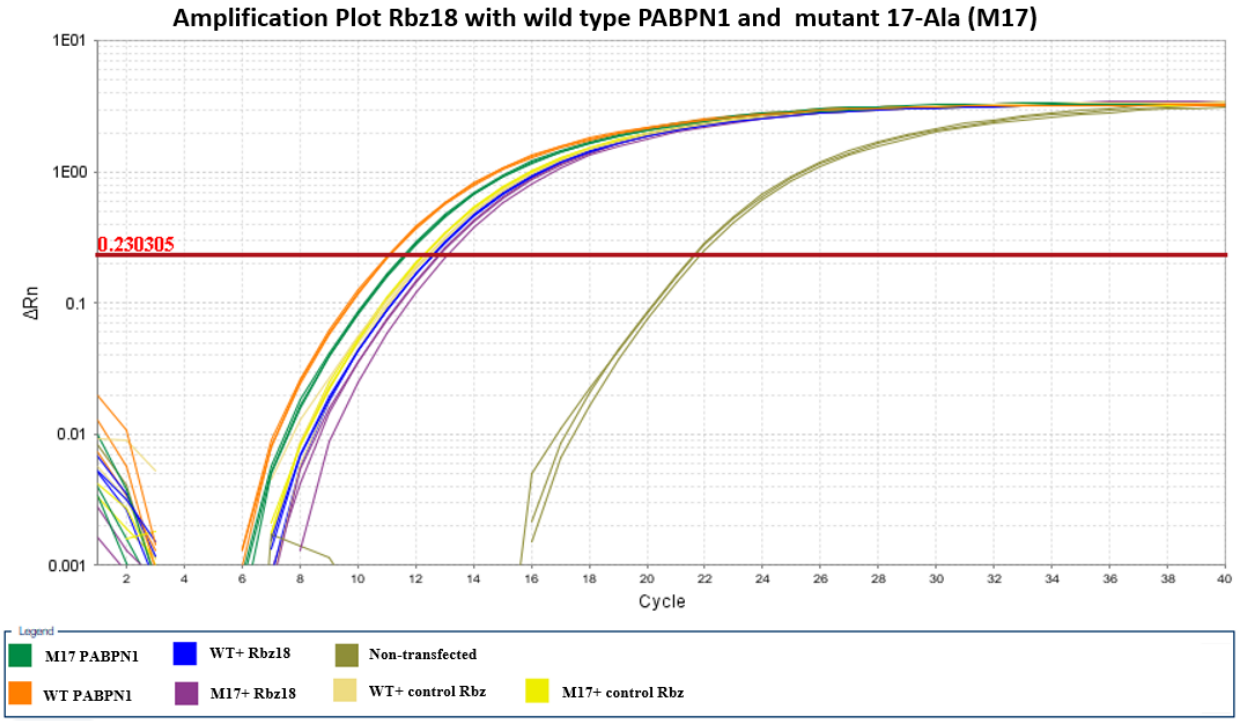

Amplification Plot Rbz13 (second try)
1E01
1E00
0.1
0.01
0.001
ΔRn
0.214941
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz13
M17 PABPN1
M17+ Rbz13
M13+ control Rbz
Non-transfected
WT PABPN1
WT+ Rbz13
M17+ control Rbz
WT+ control Rbz

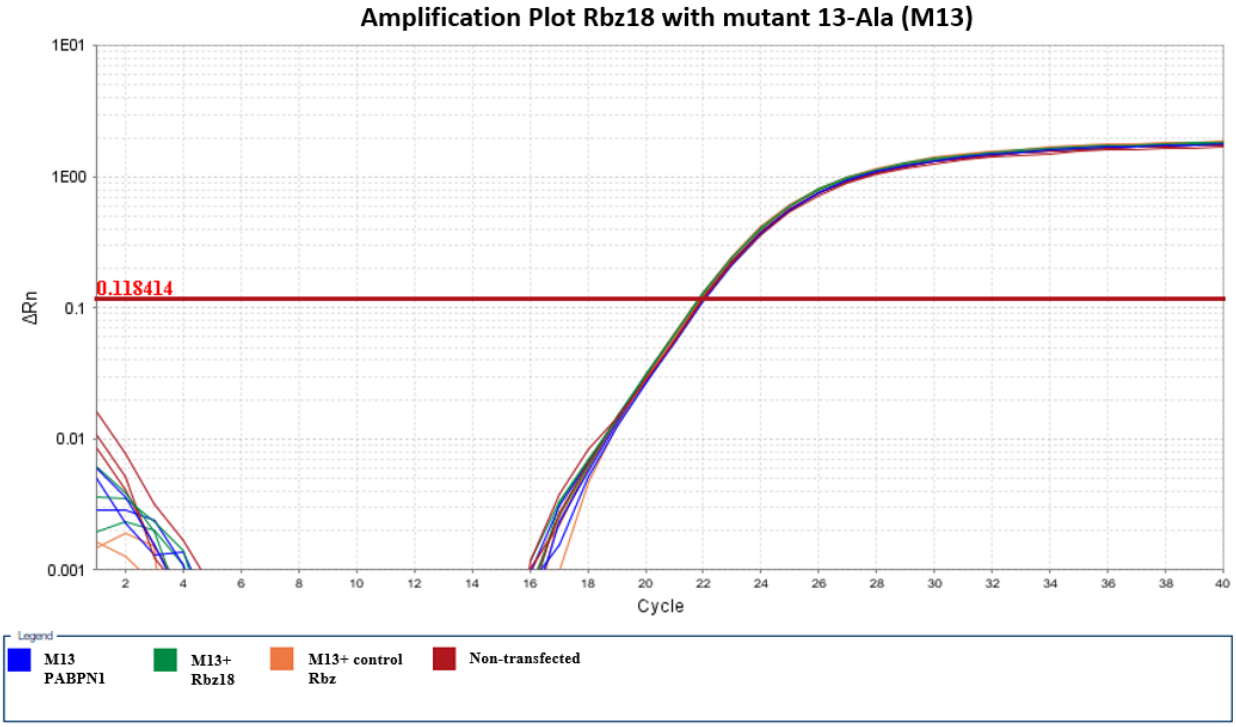

Amplification Plot Rbz14
1E01
1E00
0.1
0.01
0.001
ΔRn
0.103704
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
Cycle
Legend
M13 PABPN1
M13+ Rbz14
M17 PABPN1
M17+ Rbz14
Non-transfected
Rbz14
WT PABPN1
WT+ Rbz14

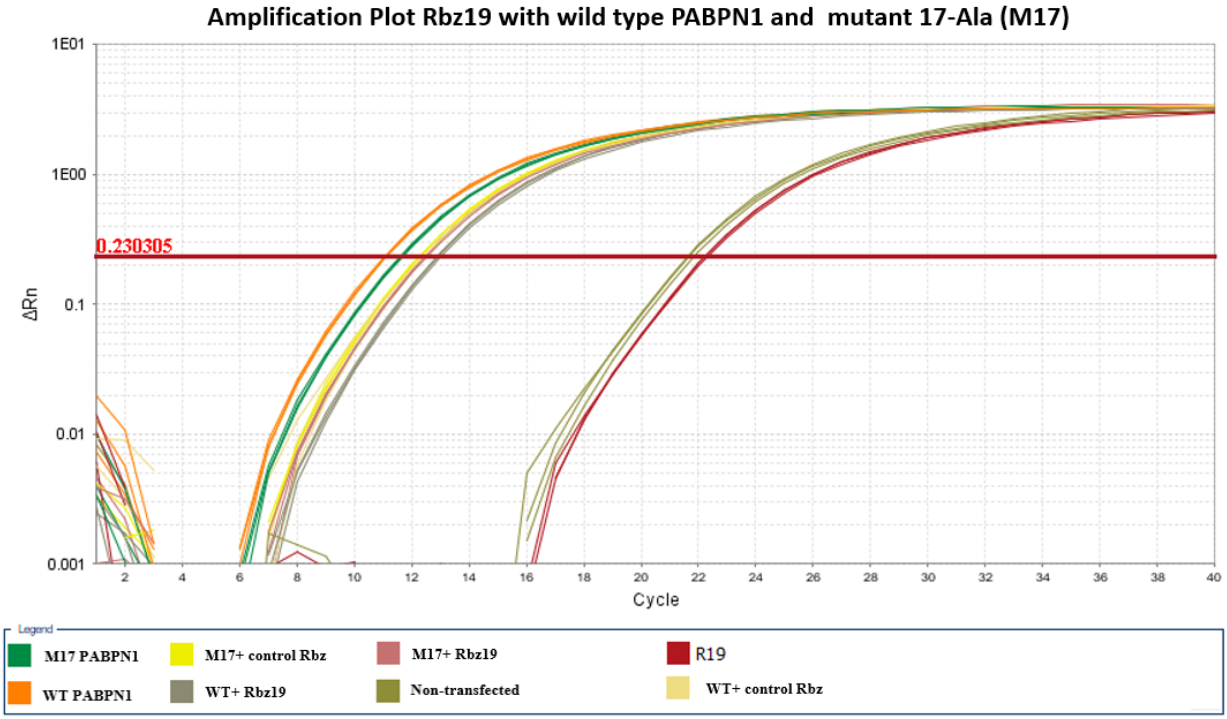

Amplification Plot Rbz15
1E01
1E00
0.1
0.01
0.001
ΔRn
0.184374
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38
Cycle
Legend
M13 + control Rbz
M13 PABPN1
M13+ Rbz15
M17 PABPN1
WT + control Rbz
M17+ Rbz15
Non transfected
WT PABPN1
M17 + control Rbz
WT+ Rbz15

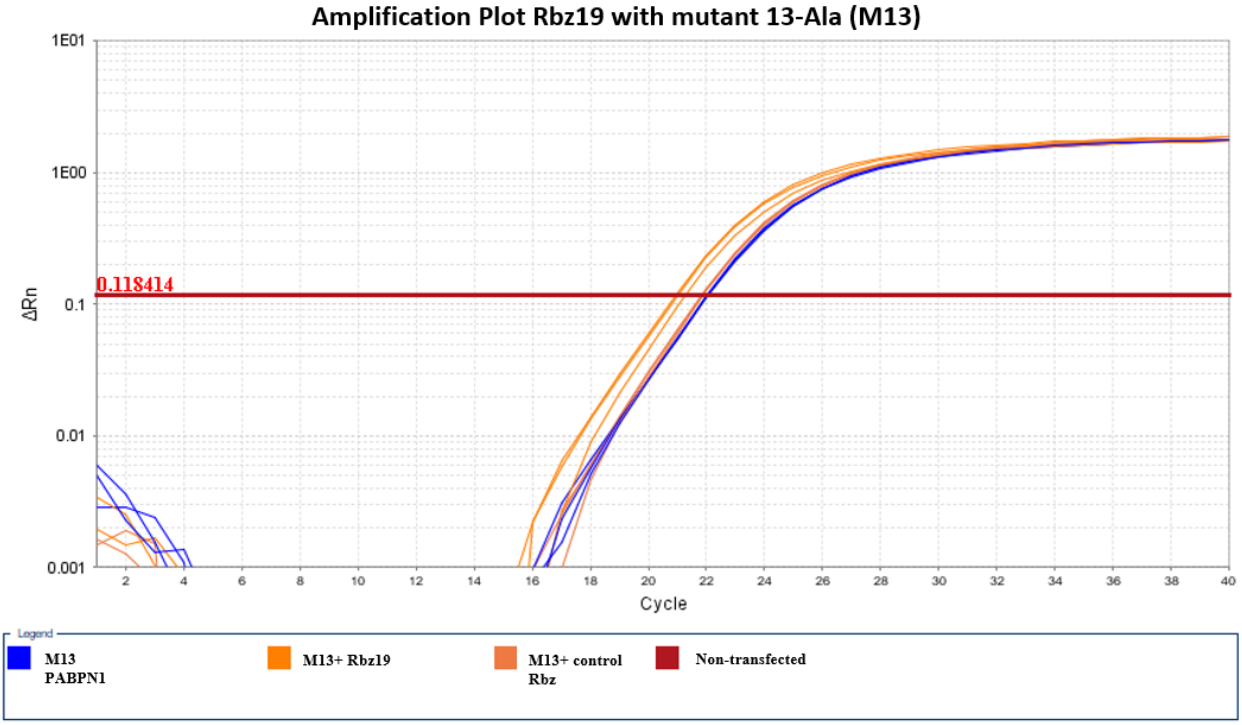

Amplification Plot Rbz16
1E01
1E00
0.184374
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 + control Rbz
M13 PABPN1
M13+ Rbz16
M17 PABPN1
WT + control Rbz
M17+ Rbz16
Non transfected
WT PABPN1
M17 + control Rbz
WT+ Rbz16

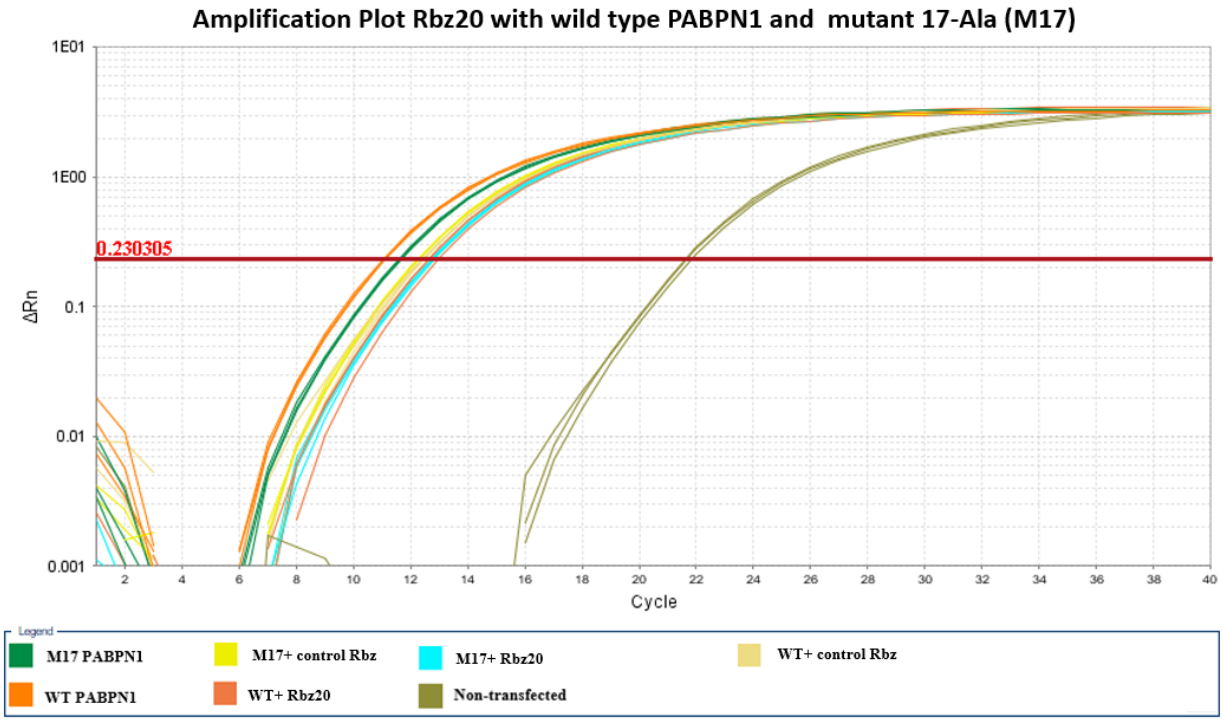

Amplification Plot Rbz17
1E01
1E00
0.184374
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 + control Rbz
M13 PABPN1
M13+ Rbz17
M17 PABPN1
WT + control Rbz
M17+ Rbz17
Non transfected
WT PABPN1
M17 + control Rbz
WT+ Rbz17


Amplification Plot Rbz18 with wild type PABPN1 and mutant 17-Ala (M17)
1E01
1E00
0.230305
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
WT+ Rbz18
Non-transfected
WT PABPN1
M17+ Rbz18
WT+ control Rbz
M17+ control Rbz


Amplification Plot Rbz18 with mutant 13-Ala (M13)
1E01
1E00
0.118414
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz18
M13+ control Rbz
Non-transfected


Amplification Plot Rbz19 with wild type PABPN1 and mutant 17-Ala (M17)
1E01
1E00
0.230305
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
M17+ control Rbz
M17+ Rbz19
R19
WT PABPN1
WT+ Rbz19
Non-transfected
WT+ control Rbz


Amplification Plot Rbz19 with mutant 13-Ala (M13)
1E01
1E00
0.118414
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz19
M13+ control Rbz
Non-transfected


Amplification Plot Rbz20 with wild type PABPN1 and mutant 17-Ala (M17)
1E01
1E00
0.230305
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M17 PABPN1
M17+ control Rbz
M17+ Rbz20
WT+ control Rbz
WT PABPN1
WT+ Rbz20
Non-transfected

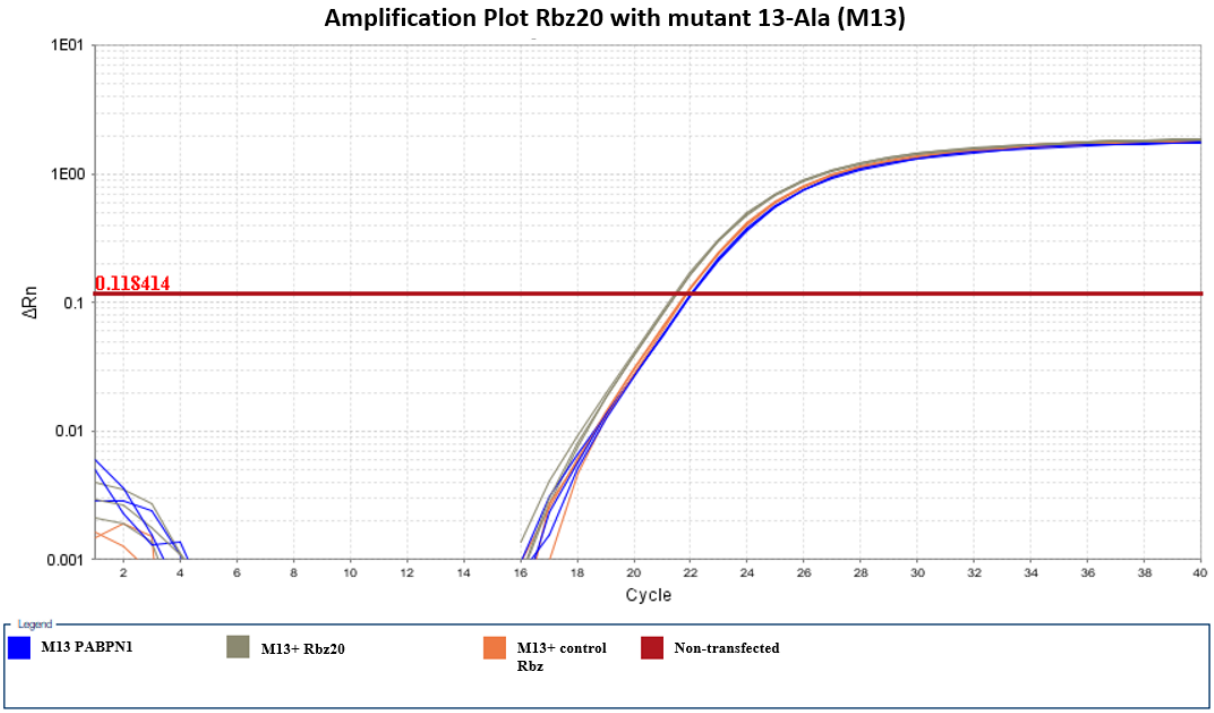

Amplification Plot Rbz20 with mutant 13-Ala (M13)
1E01
1E00
0.118414
0.1
0.01
0.001
ΔRn
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Cycle
Legend
M13 PABPN1
M13+ Rbz20
M13+ control Rbz
Non-transfected

Figure 14 RT-qPCR amplification plots for all ribozymes tested

The amplification plots for 21 ribozymes (up to Rbz 20) are presented in this figure. Every plot is titled separately and contains a legend that shows the sample content for each curve. The RT-qPCR experiment was repeated for many ribozymes in this project and plots of all trials (except the ones that contained errors) are included.

A6. Real time qPCR standard error calculation tables

A

Samples	Sample ΔCT SE	Control ΔCT SE	ΔΔCT SE
WT + Rbz	0.058	0.083	0.101257099
WT + Rbz0	0.026	0.39	0.390865706
WT + Rbz1	0.061	0.39	0.394741688
WT + Rbz2	0.125	0.39	0.409542428
WT + Rbz5	0.035	0.05	0.061032778
WT + Rbz6	0.019	0.043	0.047010637
WT + Rbz7	0.103	0.043	0.111615411
WT + Rbz8	0.026	0.062	0.067230945
WT + Rbz9	0.02	0.213	0.213936907
WT + Rbz10	0.006	0.213	0.21308449
WT + Rbz11	0.028	0.213	0.214832493
WT + Rbz12	0.081	0.317	0.327184963
WT + Rbz13	0.053	0.084	0.099322706
WT + Rbz14	0.027	0.317	0.318147764
WT + Rbz15	0.017	0.016	0.023345235
WT + Rbz16	0.053	0.016	0.055362442
WT + Rbz17	0.084	0.016	0.085510233
WT + Rbz18	0.054	0.03	0.061773781
WT + Rbz19	0.076	0.03	0.081706793
WT + Rbz20	0.125	0.03	0.128549601

B

Samples	Sample ΔCT SE	Control ΔCT SE	ΔΔCT SE
M13 + Rbz	0.068	0.07	0.09759098
M13 + Rbz0	0.11	0.052	0.12167169
M13 + Rbz1	0.025	0.052	0.05769749
M13 + Rbz2	0.115	0.052	0.12621014
M13 + Rbz5	0.114	0.029	0.11763078
M13 + Rbz6	0.085	0.081	0.1174138
M13 + Rbz7	0.14	0.081	0.16174362
M13 + Rbz8	0.051	0.028	0.05818075
M13 + Rbz9	0.078	0.12	0.14312233
M13 + Rbz10	0.28	0.12	0.30463092
M13 + Rbz11	0.04	0.12	0.12649111
M13 + Rbz12	0.01	0.146	0.14634207
M13 + Rbz13	0.061	0.096	0.11374093
M13 + Rbz14	0.11	0.146	0.18280044
M13 + Rbz15	0.051	0.039	0.0642028
M13 + Rbz16	0.023	0.039	0.04527693
M13 + Rbz17	0.023	0.039	0.04527693
M13 + Rbz18	0.035	0.046	0.05780138
M13 + Rbz19	0.102	0.046	0.11189281
M13 + Rbz20	0.06	0.046	0.07560423

C

Samples	Sample ΔCT SE	Control ΔCT SE	ΔΔCT SE
M17 + Rbz	0.029	0.02	0.03522783
M17 + Rbz0	0.007	0.02	0.02118962
M17 + Rbz1	0.008	0.02	0.021540659
M17 + Rbz2	0.153	0.02	0.154301653
M17 + Rbz5	0.02	0.26	0.260768096
M17 + Rbz6	0.022	0.022	0.031112698
M17 + Rbz7	0.418	0.022	0.418578547
M17 + Rbz8	0.026	0.053	0.059033889
M17 + Rbz9	0.032	0.012	0.034176015
M17 + Rbz10	0.018	0.012	0.021633308
M17 + Rbz11	0.053	0.012	0.054341513
M17 + Rbz12	0.1	0.042	0.108461975
M17 + Rbz13	0.024	0.042	0.048373546
M17 + Rbz14	0.059	0.042	0.072422372
M17 + Rbz15	0.009	0.026	0.027513633
M17 + Rbz16	0.207	0.026	0.20862646
M17 + Rbz17	0.022	0.026	0.034058773
M17 + Rbz18	0.094	0.022	0.096540147
M17 + Rb19	0.029	0.022	0.036400549
M17 + Rbz20	0.038	0.022	0.043908997

Figure 15. RT-qPCR standard error calculations

Every sample used in RT-qPCR experiments has a control sample with the probe for RNAPOL II to which the expression level of the target gene (PABPN1) is normalized by the qPCR machine software (QuantStudio^{TM} Real-Time PCR software v1.3). These endogenous controls are different from the controls that are used to look at the effect of the ribozyme on the level of the PABPN1 gene transcript. The control sample used for this purpose were the target samples that contained the gene without the ribozyme or the gene with a control ribozyme (the control ribozymes cannot target or affect the PABPNN1 transcripts). The mean RQ values for each transcript type (WT, M13, or M17) for every sample was then obtained by setting the control expression value to 1 and measuring the sample value based on that. This means that the effect of

a ribozyme on each of the WT, M13 or M17 transcripts is looked at independent of the other two. These values can then represent the relative expression level of that gene in presence of a certain ribozyme. The tables A, B, and C shown in this figure contain the calculations for the standard errors used to show the error bars on the comparative RT-qPCR graph for wild type, M13 and M17 respectively. The standard error values were calculated using the ΔCT SE values for every sample and its corresponding control obtained from the qPCR machine. The ΔΔCT standard error was then calculated using the following formula:

$$\sqrt{(\Delta\text{CT SE})^2{}_{sample} + (\Delta\text{CT SE})^2{}_{control}}$$

A7. Graph of comparative RT-qPCR results

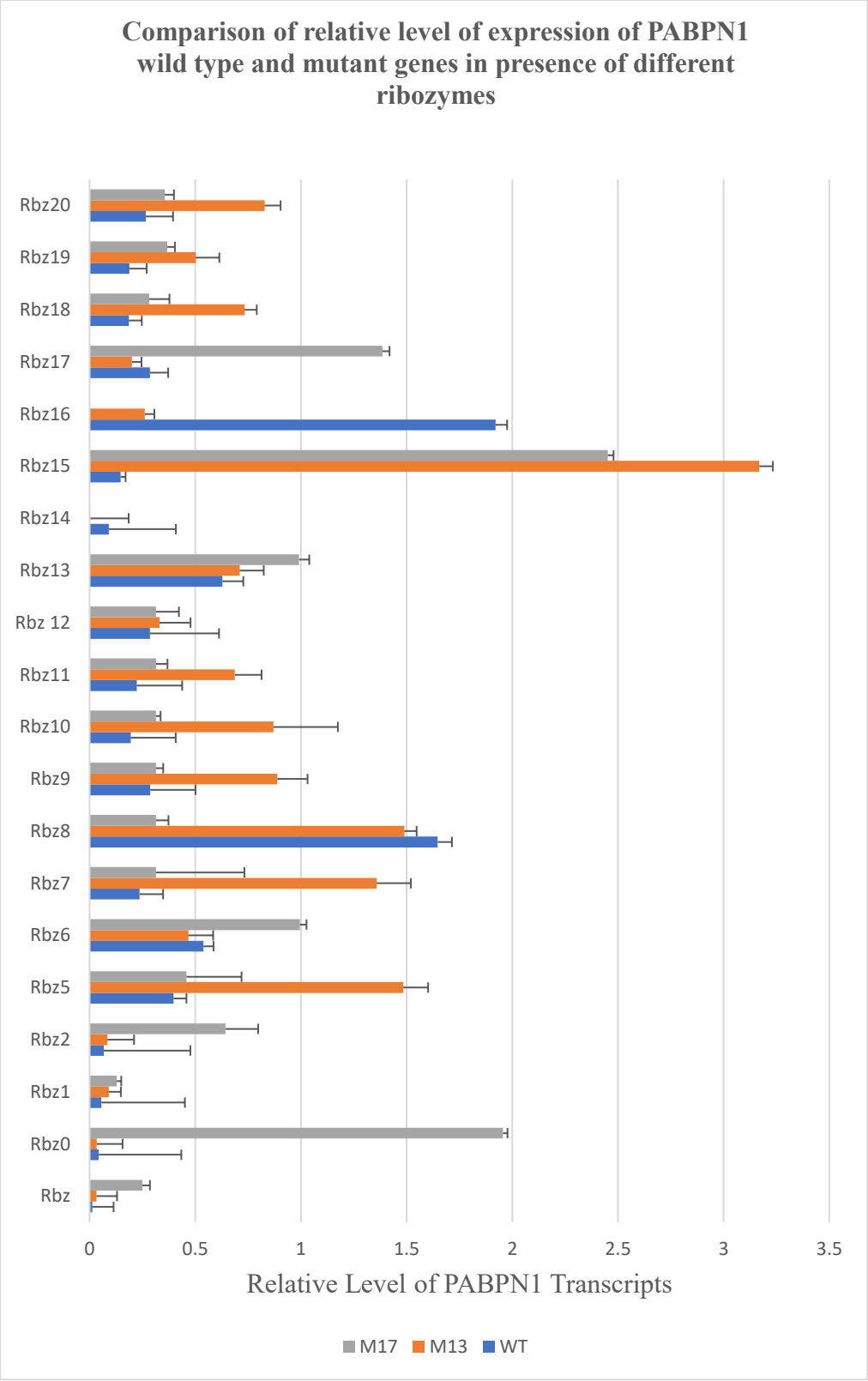


Figure 16. PABPN1 relative mRNA levels in presence of 20 ribozymes

Blue represents the wild type PABPN1, orange the mutant with 13 alanine and gray the mutant with 17 alanine residues. The results are not in line with the blot results and cannot be analyzed due to discrepancies. As depicted, the relative transcript level of Wild type PABPPN1 appears to be lower than the two mutants in presence of most of the ribozymes which cannot be explained considering the protein blot results. The issues with these results are elaborated upon in the discussion section. The RT-qPCR was repeated for many of the ribozymes such as Rbz5, Rbz8, Rbz6, Rbz13 but the results did not improve. The expression level of PABPN1 mutant 17 in presence of Rbz14 was too high compared to its control and was not presented in the graph. The ribozymes for which the expressions levels are too high are not included in the graph. The percent expression based on which this graph was generated can be found in table 4 (Appendix)

A8. Summary of RT-qPCR results table

Table 4. RT-qPCR quantitation of PABPN1 transcription in presence of ribozymes

These results are not used to categorize the ribozymes and the reasons are presented in the discussion section. qPCR and RT-qPCR are used interchangeably in the table.

Ribozyme ID	Number of times the experiment was repeated	Percent expressions when the level of expression level of the control without the ribozyme is set to 100%	Comments
Rbz	2 qPCR 2 western blots	WT+Rbz0= 1.2% M13+Rbz0 = 3.2% M17+Rbz0 = 25%	Based on qPCR results this ribozyme seems to cleave wild type, mutant 13 and mutant 17
Rbz0	1 qPCR 2 western blots	WT+Rbz0 = 4.3% M13+Rbz0 = 3.4% M17+R0 = 195.6%	Based on qPCR results this ribozyme seems to cleave wild type and mutant 13
Rbz1	1 qPCR 2 western blots	WT+Rbz1 = 5.6% M13+Rbz1 = 9% M17+Rbz1 = 12.8%	Based on qPCR results this ribozyme seems to cleave wild type, mutant 13 and mutant 17

Rbz2	1 qPCR 2 western blots	WT+Rbz2 = 6.7% M13+Rbz2 = 8.4% M17+Rbz2 = 64.3%	Based on qPCR results this ribozyme seems to cleave wild type, mutant 13 and mutant 17
Rbz3	1 qPCR 2 western blots	WT+Rbz3 = 163% M13+Rbz3 = 75% M17+Rbz3 = 74.6%	Based on qPCR results this ribozyme seems selective only for mutant transcripts, but the blot results do not confirm selectivity.
Rbz4	1 qPCR 2 western blots	WT+Rbz4 = 100% M13+Rbz4 = 17% M17+Rbz4 = 124.5%	Based on qPCR results this ribozyme seems to be selective for M13, but the blot results do not confirm this.
Rbz5	2 qPCR 2 western blots	WT+Rbz5 = 317% M13+Rbz5 = 36.3% M17+Rbz5 = 64.2%	Based on qPCR results this ribozyme seems to cleave mutant 17 and mutant 13 and the western blot results confirm this.
Rbz6	2 qPCR 2 western blots	WT+Rbz6 = 53.9% M13+Rbz6 = 46.7% M17+Rbz6 = 31.4%	Based on qPCR results this ribozyme seems to cleave wild type, mutant 13 and mutant 17
Rbz7	1 qPCR 2 western blots	WT+Rbz7 = 23.6% M13+Rbz7 = 135.8% M17+Rbz7 = 41.3%	Based on qPCR results this ribozyme seems to cleave the wild type, and mutant 17
Rbz8	2 qPCR 2 western blots	WT+Rbz8 = 164.7% M13+Rbz8 = 1.489 M17+Rbz8 = 29.2%	based on qPCR this ribozyme seems to cleave both mutant transcripts and the blot results confirm this.
Rbz9	1 qPCR 2 western blots	WT+Rbz9 = 28.7% M13+Rbz9 = 88.8% M17+Rbz9 = 43.7%	Based on qPCR results this ribozyme cleaves WT, m13, m17.

Rbz10	1 qPCR 2 western blots	WT+Rbz10 = 19.4% M13+ Rbz10 = 87% M17+R10 = 41.2%	Based on qPCR results this ribozyme cleaves WT, m13, m17.
Rbz11	1 qPCR 2 western blots	WT+Rbz11= 22.3% M13+Rbz11 = 68.7% M17+Rbz11 = 62.5%	Based on qPCR results this ribozyme cleaves WT, m13, m17.
Rbz12	1 qPCR 2 western blots	WT+Rbz12= 28.5% M13+Rbz12 = 33.1% M17+Rbz12 = 81.8%	Based on qPCR results this ribozyme cleaves WT, m13, m17.
Rbz13	2 qPCR 2 western blots	WT+Rbz13= 62.8% M13+Rbz13 = 71% M17+Rbz13 = 99.1%	Based on qPCR results this ribozyme cleaves WT, m13.
Rbz14	1 qPCR 2 western blots	WT+Rbz14= 9% M13+Rbz14 = 0.2% M17+Rbz14 = 524% (not included in graph)	Based on qPCR results this ribozyme cleaves WT, m13 results cannot be analyzed due to spuriously high amplification of M17 transcript in presence of the ribozyme.
Rbz15	1 qPCR 2 western blots	WT+Rbz15= 14.6% M13+Rbz15 = 316.8% M17+R15 =245.1%	results cannot be analyzed due to spuriously high amplification of transcripts.
Rbz16	1 qPCR 2 western blots	WT+Rbz16= 192% M13+Rbz16 = 26.1% M17+Rbz16 =861% (not included in graph)	results cannot be analyzed due to spuriously high amplification of transcripts.
Rbz17	1 qPCR 2 western blots	WT+Rbz17= 28.6% M13+Rbz17 = 20% M17+Rbz17 =138.5%	Based on qPCR results this ribozyme cleaves WT, M13
Rbz18	1 qPCR 2 western blots	WT+Rbz18 = 18.5% M13+Rbz18 = 73.3% M17+Rbz18 = 28.1%	Based on qPCR results this ribozyme cleaves M17, wild type and M13.
Rbz19	1 qPCR 2 western blots	WT+Rbz19 = 18.8% M13+Rbz19 = 50.2% M17+Rbz19 = 36.7%	Based on qPCR results this ribozyme cleaves M17, wild type and M13.

Rbz20	1 qPCR 2 western blots	WT+Rbz20 = 26.6% M13+Rbz20 = 82.8% M17+Rbz20 = 35.5%	Based on qPCR results this ribozyme cleaves M17, wild type and M13.
Rbz2840	2 western blots	qPCR was not done for these samples because blot results did not indicate significant reduction in protein levels	No significant reduction in protein level based on blot results
Rbz2841			
Rbz2842			
Rbz2843			
Rbz2844			
Rbz2845			
Rbz2846			
Rbz2847			